Positional Accuracy Improvement - A Necessary Tool for Updating and Integration of GIS Data

Frank GIELSDORF, Lothar GRUENDIG and Bernd ASCHOFF, Germany

Key words: Geo-Referencing, GIS, Positional Update

SUMMARY

Common geographic information systems use spatial data with a point accuracy of about 0,5 down to 5 meter. The geometrical quality of the spatial base data is improved stepwise by integration of precise geodesic measurements as well as by air photos or existing field book measurements. The mounted spatial thematic data should profit from this accuracy improvement without loosing its internal geometrical quality. But the most GIS data models do not support such strategies, and update transformations lead to inconsistency because proximity fitting principles are neglected.

The paper shows how the requirements concerning PAI can be provided using adjustment techniques. A data maintaining strategy is presented which regards point coordinates just as a view on redundant primary data. PAI updates are to be seen as the generation of a new view applying adjustment techniques. The necessity of a point topology in the GIS structure is pointed out.

ZUSAMMENFASSUNG

Zurzeit verwenden Geoinformationssysteme Geodaten, die nur eine Genauigkeit von 0,5 bis 5 Meter haben. Die Präzision der Geo-Basisdaten wird durch die Integration von geodätischen Messungen, Luftbildern oder vorhandenen Vermessungszahlen schrittweise verbessert. Davon sollen die angeschlossenen Geo-Fachdaten profitieren, ohne ihre innere geometrische Qualität zu verlieren. Doch von den meisten GIS Datenmodellen wird ein solches Vorgehen nicht unterstützt, es kommt bei nachträglichen Transformationen zu Inkonsistenzen, weil Prinzipien zur Wahrung der Nachbarschaft nicht beachtet werden.

Die Abhandlung zeigt, wie die Anforderungen bezüglich PAI durch Ausgleichungsalgorithmen effizient unterstützt werden können. Es wird eine Strategie der Datenverwaltung vorgestellt, welche die absoluten Punktkoordinaten als Sicht auf redundante Primärdaten betrachtet. Die Notwendigkeit einer Punkt-Topologie in der GIS Struktur wird unterstrichen.

Positional Accuracy Improvement - A Necessary Tool for Updating and Integration of GIS data

Frank GIELSDORF, Lothar GRUENDIG and Bernd ASCHOFF, Germany

1. INTRODUCTION

1.1 Initial Situation

The initial data acquisition of GIS is completed in many cases. Foundations of the point coordinates are mainly analogue maps which were digitized and geo-referenced. The resulting geometrical quality of GIS reflects the quality of the underlying maps. The point accuracy commonly encompasses a range from 0.5m down to 5m. Present spatial base data are the geo-reference for spatial thematic data and descriptive data as facility lines, addresses, etc. Thematic objects were constructively attached to base objects using CAD tools whereby geometrical redundancy remained unconsidered. Available measures from field books exist just as character attributes.

1.2 Appearance and Problems of the PAI Process

Geodesic and photogrammetric observations (GPS, tachometric measurements, air photos) are the product of surveyor's daily work. They result in better point coordinates and are provided to improve spatial base data in GIS. However, this integration is done stepwise.

-As point positions remain unchanged in reality, update coordinates lead to a virtual displacement of the related points in the GIS. Of course, without any consideration of neighborhood relationships to other points at all, the relative geometry between updated and unchanged points in the GIS would be highly violated (e.g. a facility line offset to a Building is now 5.5 meter instead of earlier 2 meters, or the facility line is even crossing the Building). Of course, such a coordinate exchange would never be accepted.

2. DATA MANAGEMENT

2.1 Some Reasons of PAI Problems

First of all it has to be made sure that PAI proceeds can be realized because of the data concept. The history of geo information has proven that coordinates updates were not designed as a matter of course.

GIS are in general designed to consider geometrical parameters (coordinates) as deterministic values. This architecture reflects the view of a software engineer, which can be described as follows:

‘If a point does not change in reality, it has not to be changed in the database. Each relative geometry measure (distance, angle etc) can be calculated from coordinates. Coordinates and relative measures are completely equivalent.

On the other hand, the view of a surveyor is different:

‘Coordinates are always calculated from observations. Observations are redundant aleatory values. From this it follows that coordinates are aleatory values as well and over that correlated. The accuracy of relative geometry is higher than the absolute accuracy. Coordinates and relative measures are therefore **not** equivalent. The derivation of coordinates from observations is unique but not reversible.’

2.2 Integration of Relative Geometry

At a virtual displacement of points the different accuracy of relative and absolute geometry has to be considered. This would be theoretically possible by introduction of very large covariance matrices, but this is not practicable. Alternatively relative measures are introduced as carrier of relative accuracy information. Two general cases can be discerned:

Case 1: Original observations do not exist

A typical instance are digitized map coordinates. A common Helmert or affine transformation does not consider the correlations between neighboring points, therefore artificial observations, based on substantiated hypotheses, are generated and introduced in an adjustment. Commonly the neighboring information is derived from a delaunay triangulation whose triangles or triangle sites are used as carrier of these artificial observations. As observation types are distances or coordinate differences frequently used. But also geometrical constraints like co-linearity or orthogonallity can be seen as artificial observations, based on the hypothesis that buildings are commonly rectangular respectively points on one borderline are co-linear.

Case 2: Original observations exist

Original observations can result from available field books, newer on site measurements or aerial flights, observation types are distances, directions, local coordinates etc.

Because of their redundancy, the consideration of relative geometry measures leads to an adjustment problem in any case. Furthermore, only adjustment techniques provide the option to integrate artificial and original observations to determine unique absolute coordinates.

It is known that the best consideration of neighborhood relationships is warranted using proximity fitting adjustment methods where artificial observations between points are integrated. Advanced adjustment programs use for that task finite element methods based on triangles. Nevertheless, the real observations can completely be introduced in these proximity fitting adjustment processes. The adjustment program SYSTRA provides these options.

2.3 The Role of Topology

By definition topology is a system of subsets defined on a set. A graph is a special type of topology which consists of bivalent subsets (edges) defined on a set of nodes. In this paper topology should be interpreted in narrow sense as ‘rubber geometry’. Points can be displaced, but their fundamental neighborhood relationships do not change (e.g. a tree should not slide into a building because of any transformation). In consequence this means that topology is invariant against transformations of geometry.

The entity type point is the link between relative and absolute geometry. However, in most GIS databases it does not exist as an independent entity type. Points are either part of a higher-level object (shape) or they are identified by their coordinates. According to the principles of programming and database theory the identifier of an object remains constant during its live cycle. If the identifier changes it is conterminously with the disappearance of the object. But even this happens during a PAI process.

PAI can be seen as a long transaction on a database. This means that a partition of the database will be checked out, processed by an external program and checked in again. The affected partition remains locked in the database during transaction. PAI has, like any transaction, to follow the so called ACID principles what means atomicity, consistency, isolation and durability. Therefore it is necessary for PAI to give each point, if necessary temporarily, a unique identifier which is independent of its coordinates.

PAI requires a strict separation of topological and geometrical information in the data model. Leading GIS software providers have recognized this problem and started to create extended data models. Figure 1 illustrates the role of a topological layer in the data model with respect to PAI.

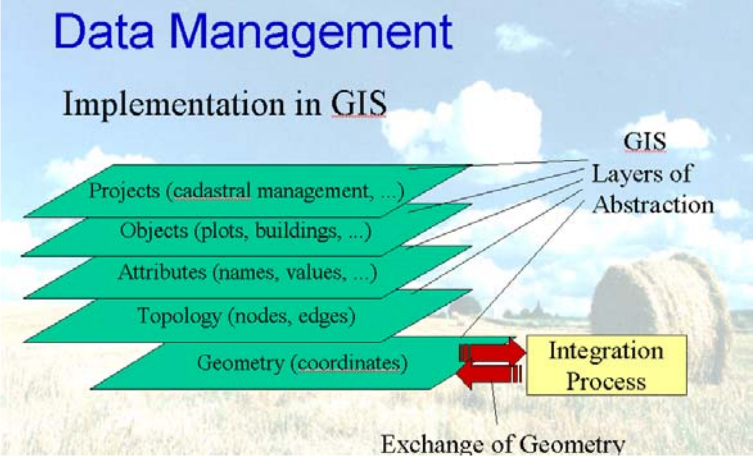

Figure 1: Separation of Topology and Geometry

2.4 Modeling of Point Identities

As it was shown, the input data of a PAI process are descended from different data sources. In order to integrate these different data sets it is necessary to detect point identities in between them. In the majority of cases points will have either different or no identifiers anyway. The determination of point identities can be in practice one of the most time consuming parts of the whole process. Therefore it is sensible to use sophisticated matching tools to solve this problem automatically or at least semi automatically.

There are two principles to model point identities. Either the identity information is expressed topologically or geometrically. Most common is the topological modeling of identities by giving the same point identifier to corresponding points in different data sets. But this method can lead to problems because of inevitable point confusions. Confusions affect only indirect to the residuals of corresponding observations. In order to eliminate confusions it is necessary to ‘disfuse’ points by generation of new point objects whereby referential integrity requirements have to be considered.

An alternative approach is the geometrical modeling of point identities. Instead of a common identifier an identity observation is introduced. This means that between two possibly identical points coordinate difference with the value zero is observed. The square sum of the residuals of such identity observation is χ^2-distributed and can be tested for significance. Misidentifications can then easily be detected and eliminated.

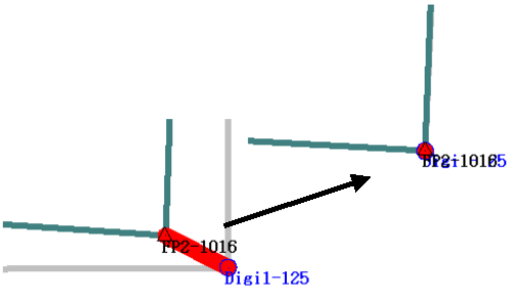

Figure 2: Point identity handling

Figure 2 shows the principle of a point identity observation. It is a relative measure between two points with different point numbers. The point identity is weighted with a standard deviation derived from connected observations (e.g. map accuracy). It reacts like a ‘rubber band’ (see left part of the figure) with the elasticity of its weight and can be analyzed like all other observations (see Data Snooping). Not reliable measures can easily be removed without violating the topology. If all remaining point identities are reliable, their standard deviation is fixed (set to zero), and the connected points get the same coordinates (see right part of the figure). Finally, the points can be melted by the GIS to get a topology free of redundancy.

3. ADVANCED ADJUSTMENT TECHNIQUES

The significance of adjustment techniques for transformation problems is recognized long ago. To apply the least squares method following C.F. Gauß is the usual geodesic practice for two dimensional transformations with redundant identities from one Cartesian system into another. This classical adjustment method can be expanded to a simultaneous transformation of multiple systems, subsequently called ‘Interconnected Transformation’.

3.1 Interconnected Transformation

Table 1 shows the comparison between the classical and the interconnected transformation approach. The difference consists in extension of the set of unknowns. In addition to the transformation parameters X_0, Y_0, a, o, the coordinates of the interconnection points will be introduced as unknowns. This approach leads to a non-linear adjustment problem.

Table 1: Comparison between the approaches

	Classic	Interconnected
Observations	x_i, y_i	x_i, y_i
Target Function	$\sum_{i=1}^{n_p} v_{x_i}^2 + v_{y_i}^2 \overset{!}{=} \min$	$\sum_{i=1}^{n_p} v_{x_i}^2 + v_{y_i}^2 \overset{!}{=} \min$
Unknowns	X_0, Y_0, a, o	$X_0, Y_0, a, o, \mathbf{X_i}, \mathbf{Y_i}$

Figure 3 shows differences between the transformations concerning point referencing and connections. In the left part six single transformations with a minimum number of reference points have to be taken. In the right part only one interconnected transformation has to run. The point identities are able to substitute reference points and to connect map borders.

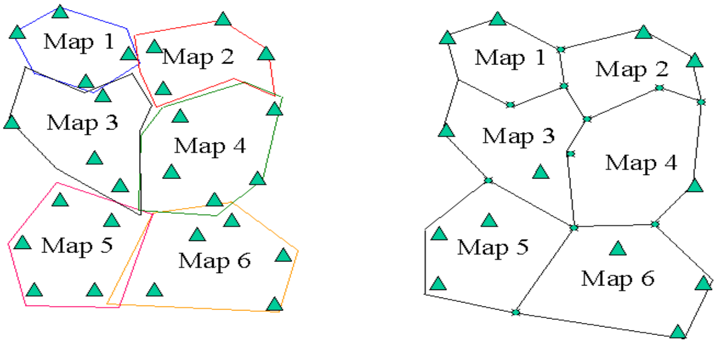

Figure 3: Transformation principles concerning point connections

The data snooping method (Baarda) is an appreciated analysis with respect to objective determination of observation blunders. The normalized residuals (NV) of observations are derived from the normal equations of the adjustment model.

$$\mathbf{x} = (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l} \qquad (0.1)$$

In classical transformations, coordinate residuals would be divided by their empirical standard deviations (a posteriori). However, because of missing observation type variation and unique accuracy, they are comparable anyway.

3.2 Proximity Fitting

Proximity fitting methods are applied to keep neighborhood relationships. They substitute a usual single system transformation. The difference can be summarized as follows:

Case 1: Usual transformation

Average transformation parameters are adjusted using redundant identity points. The number of parameters can be 3, 4, 5 or 6. With this parameter set, all uncontrolled points are calculated the same way, in so far each point is influenced by every identity point.
Case 2: Proximity fitting

The result of a usual transformation can be seen as the first step of proximity fitting. In a second step the mapping approach is extended by the introduction of relative geometry information.

An old and well known method is the distance weighted interpolation. Artificial coordinate differences between identical points, the connection points and the new points are introduced into the adjustment. These 'pseudo observations' are weighted dependent on the distances.
Figure *4* shows that there are no direct neighborhood relationships between the interpolated points. Additionally, the result depends on the number of identity points which create residuals. The method is not suitable to model direct neighborhood relationships. The resulting displacements of the new points are dependent of the density and distribution of the identical points.

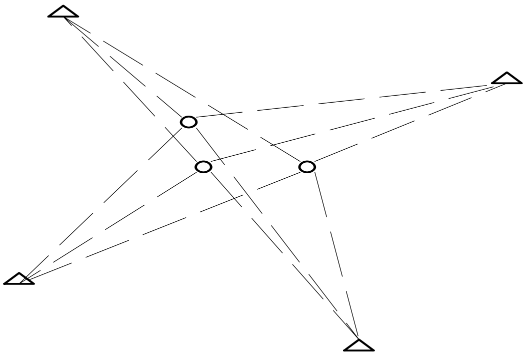

Figure 4: Interpolation using weighted distances

But the residuals in identical points can be seen as discrete representatives of an area wide acting systematic and it should be the aim of proximity fitting to model even this area systematic. Therefore advanced methods use the delaunay triangulation to model neighborhood relationships directly. The resulting displacements are here independent of the density and distribution of identity points.

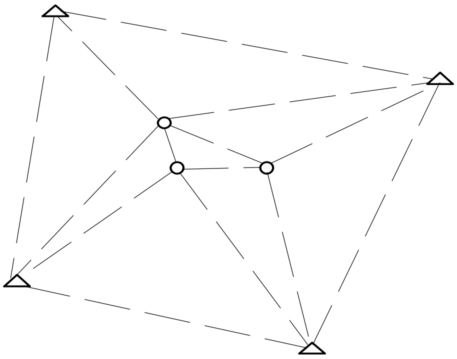

Figure 5: Membrane Method Using Delaunay-Triangulation

However, the different methods vary in their functional and stochastic modeling of the artificial observations. The here suggested approach is called ‘membrane method’ following the experience of the authors in lightweight structure modeling. This method uses as functional model coordinate differences along the triangle sites, what leads to linear residual equations with a very stabile convergence behavior. The stochastic model is derived from finite element methods, and it simulates the behavior of a rubber membrane. The sequence of operations is as follows:

- With the interconnected transformation all identity point residuals are calculated with usual transformation types (4, 5 or 6 PT).
- For each local system (mostly a map) a separate Delaunay-Triangulation is made.
- The local coordinates are substituted by coordinate difference observations along the triangle sites.
- The proximity fitting is run as an adjustment calculation.

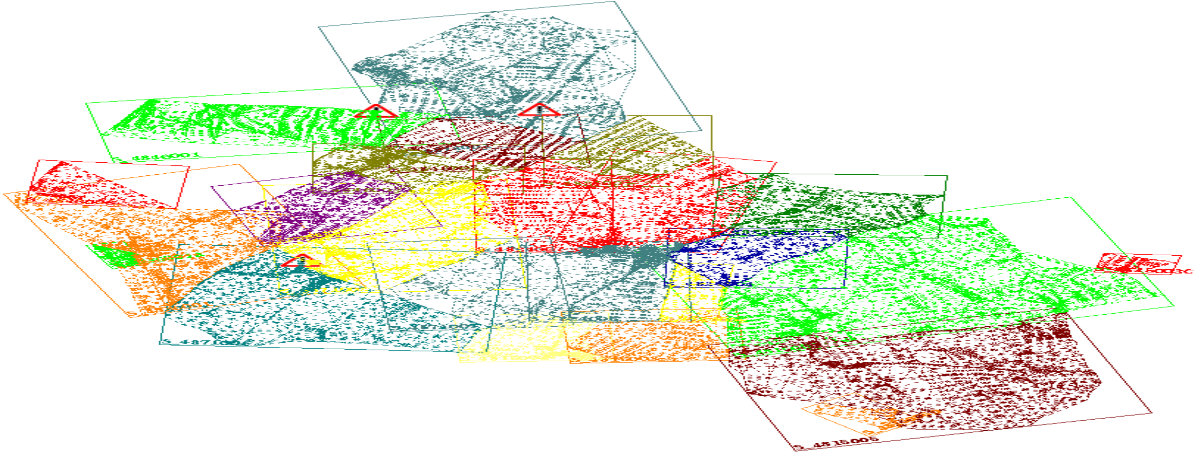

Figure 6: Separate Delaunay-Triangulations for Many Local Systems

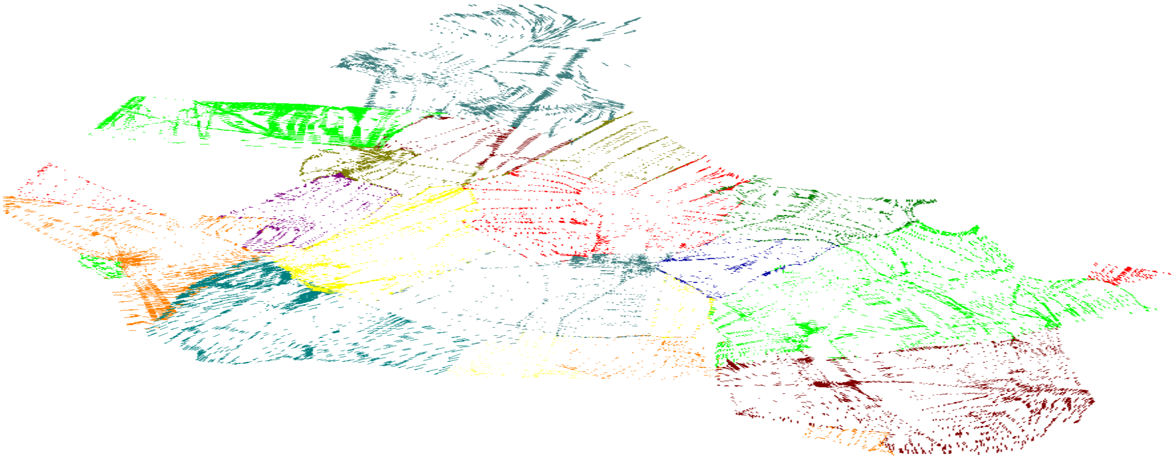

Figure 7: Displacement Vectors of Interconnected Proximity Fitting

3.3 Integration of Geodesic Measurements

A point accuracy improvement can be seen as the integration of already existing GIS coordinates, GPS and polar measurements, field book measures and geometrical constraints. This is valid for spatial basic data as well as for spatial thematic data. The following example shows a map transformation with integration of relative measures. The example contains two spatial data sets with linear constraints of the facility lines, parallels with offset between these lines and facility line offsets to buildings. The point identity observation type is used.

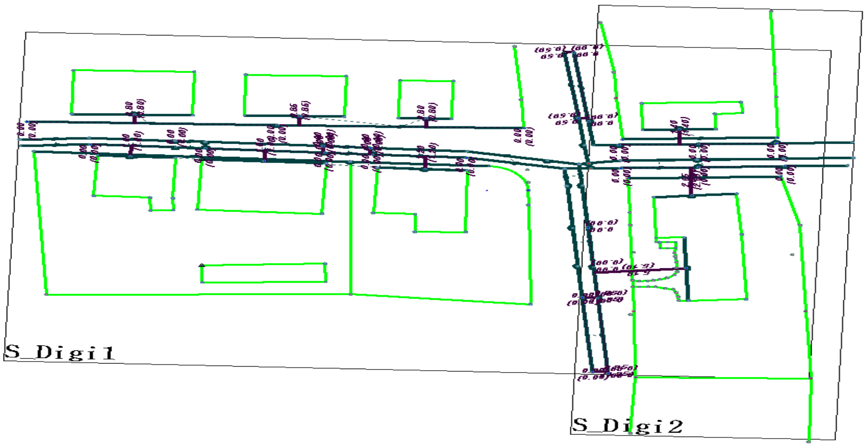

Figure 8: Spatial Data with Geometrical Constraints

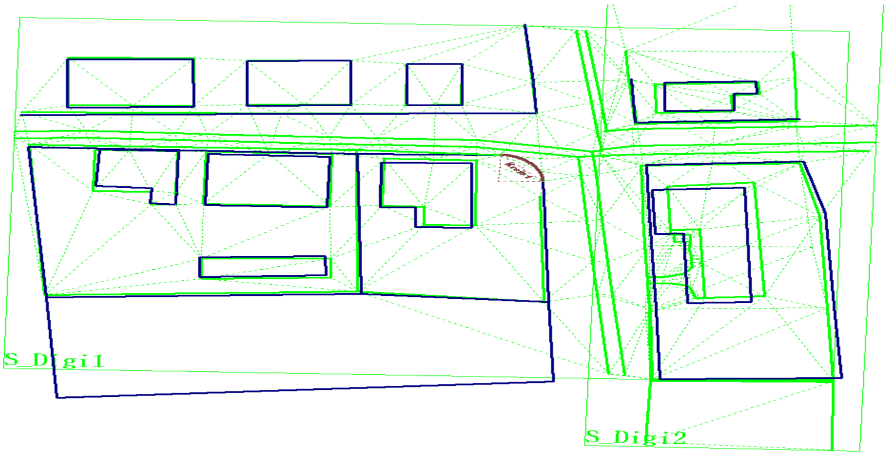

Figure 9: Spatial Thematic and new Spatial Base Data, Transformation not done as yet

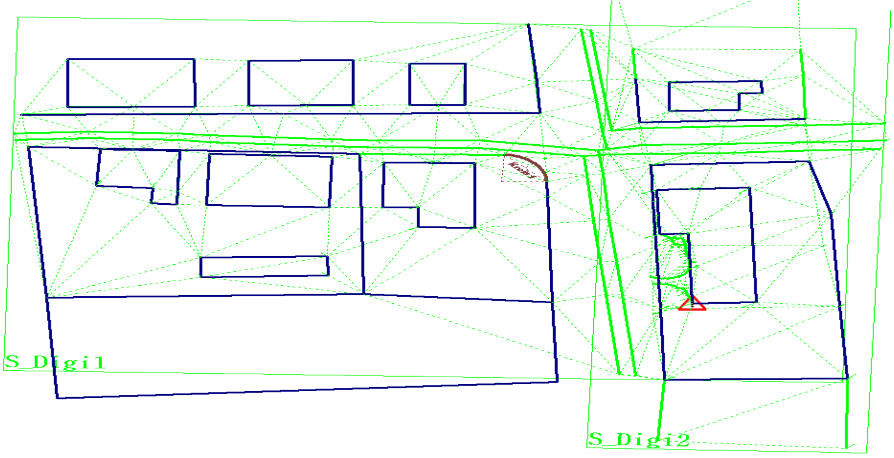

Figure 10: Proximity Fitting without Geometrical Constraints

Figure 10 shows a problem at the building right bottom. Real estate border and building edge should be co-linear.

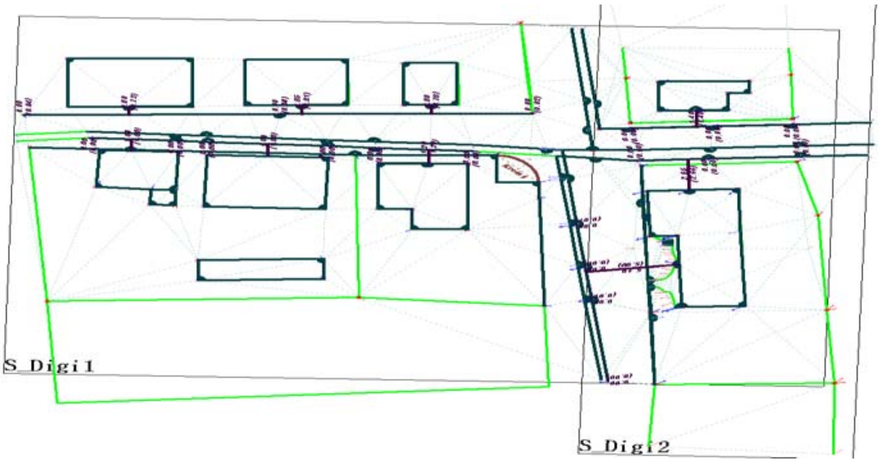

Figure 11: Proximity Fitting with Geometrical Constraints

Figure *11* shows the wanted result if relative measures are introduced. The stochastical analysis provides information about the significance of geometrical constraints.

4. PROJECT EXAMPLES

4.1 Geo-Referencing for the ALK in Germany

Many projects of first data acquisition for the German Automated Real Estate Map (ALK) in East Germany showed the efficiency of interconnected transformation to geo-reference a large number of digitized maps simultaneously. However, the problems to be solved were less of a mathematical type, as the interconnected transformation and its associated proximity fitting are already a powerful working technique as proven with SYSTRA.

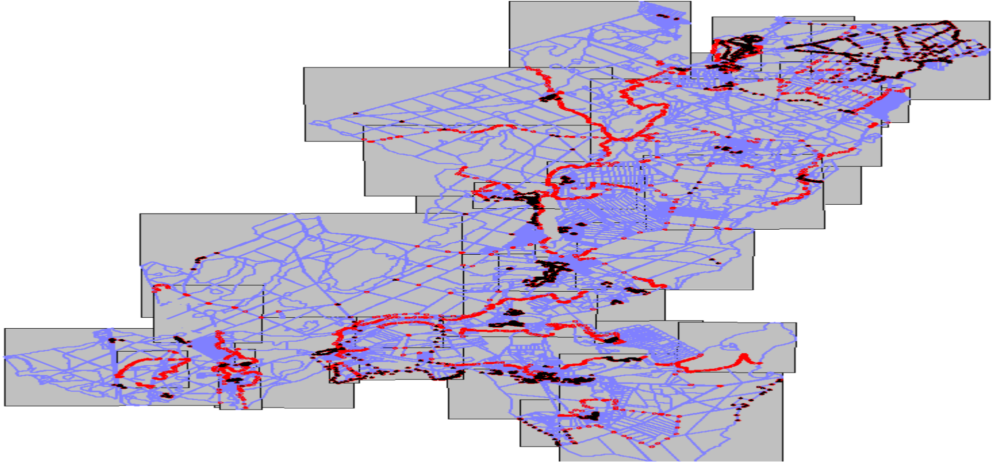

Figure 12: Interconnected Transformation of Many Insular Maps

Much more deciding was the realization of the workflow as a long transaction on a GIS database. Unique object identifiers were necessary. The map connection was realized using point identity observations. It helped get about 40% more efficiency compared to the classic connection technique during the execution of the geo-referencing project part. The automatic renumbering was done after the geo-referencing work.

However, it was realized that for certain regions with maps of bad quality geodesic measurements had to be added to the project. These observations from field books were able to give reliable map overlapping zone connections. They were suited partly to get a better relative accuracy, and if coherent with reference points they even effected an improvement of absolute accuracy.

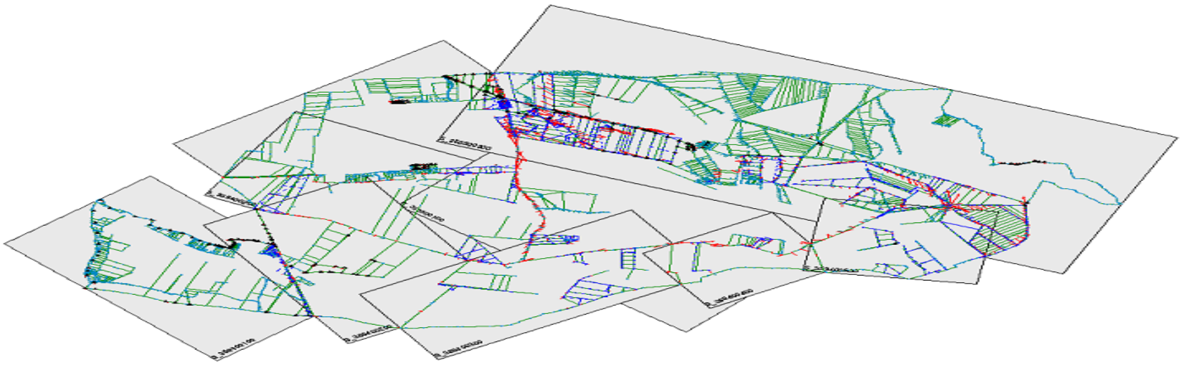

Figure 13: Simultaneous Map and Geodesic Measurement Analysis

4.2 Proximity Fitting of Local Referenced Point Fields in Hamburg

The department of geo-information und surveying in Hamburg decided to transform large point fields referencing different local frames (Gauß-Krüger and partly Soldner) into one unique reference frame (ETRS89 with UTM projection). An interconnected transformation approach was chosen for the calculation. Combined with an integrated GPS campaign, this procedure provided homogeneous and precise coordinates.

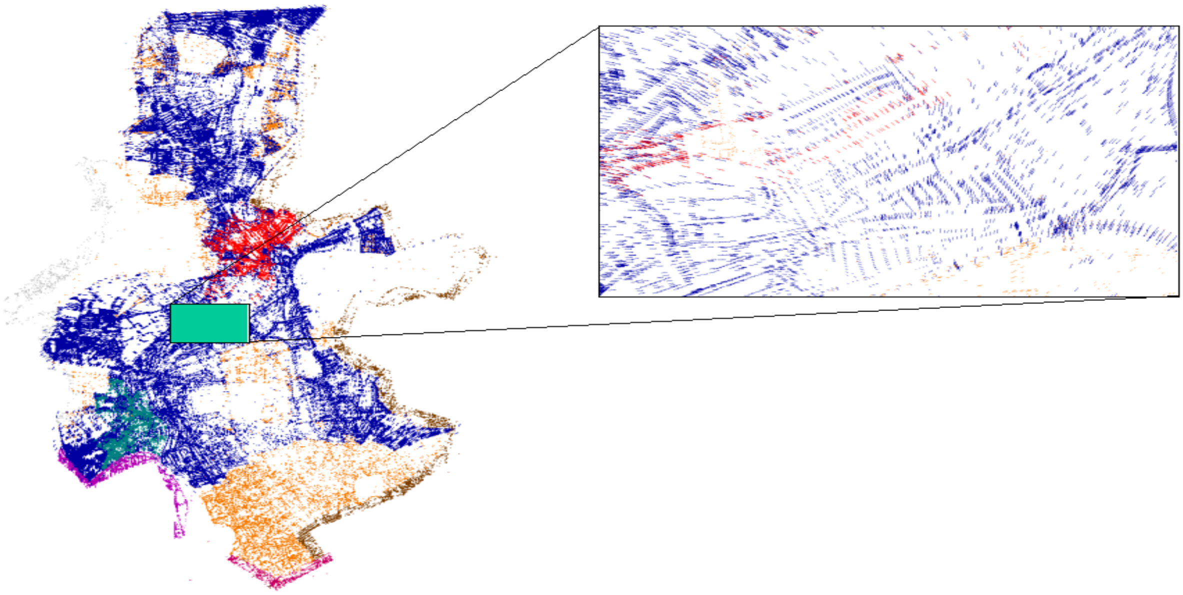

Figure 14: Proximity Fitting of Local Reference Frames

The strategy was to group point fields to blocks. Each block contained about 200.000 points. The blocks were bounded by points which were determined with GPS. The number of additional points in the blocks to be determined by GPS measures depended of the accuracy analysis result of the applied transformation. In order to keep all neighborhood relationships in the border and overlapping zones, a proximity fitting was performed after the interconnected transformation.

REFERENCES

Gründig L., Gielsdorf F., Scheu M., Dreesmann R., “Verkettete Transformation, Mustertechnologie”, Surveying Department Brandenburg, internal paper (1999)

‘Umsetzung der analogen Liegenschaftskarten in die digitale Liegenschaftskarte im ländlichen Raum”, Surveying Department Brandenburg, internal paper (1999)

„Überführung des Grenznachweises der Freien und Hansestadt Hamburg in das Bezugssystem ETRS89“, Department of Geo-Information and Surveying Hamburg, internal paper (2001)

BIOGRAPHICAL NOTES

Dr. Frank Gielsdorf, born 1960. Graduated in 1987 as Dipl.-Ing. in Surveying from Technical University of Dresden. Obtaining doctorate degree in 1997 from Technical University of Berlin. Since 1995 Assistant Professor at the Department of Geodesy and Geomatics, Technical University of Berlin.

Prof. Dr. Lothar Gruendig, born in 1944. Graduated in 1970 as Dipl.-Ing. in Surveying and obtaining doctorate degree in 1975, both from University of Stuttgart. From 1970 to 1977, work as Assistant Professor and until 1987 senior research assistant at University of Stuttgart. Scientist at Scientific Center of IBM in Heidelberg on data bases 1984-1982 and guest scientist at Calgary University for 4 months in 1983. Since 1988 Professor of Geodesy and Adjustment Techniques at the Department of Geodesy and Geomatics, Technical University of Berlin.

Bernd Aschoff, born 1962. Graduated in 1987 as Dipl.-Ing. in Surveying from Technical University of Berlin. Since 1995 managing director of the technet gmbh.

CONTACTS

Frank Gielsdorf
Technische Universität Berlin, Sekretatiat H20
Strasse des 17. Juni 135
10623 Berlin
Germany
Tel. + 49 30 314 22 374
Fax + 49 30 314 21 119
Email: gielsdorf@inge3.bv.tu-berlin.de
Web site: www.survey.tu-berlin.de

Lothar Gruendig
Technische Universität Berlin, Sekretatiat H20
Strasse des 17. Juni 135
10623 Berlin
Germany
Tel. + 49 30 314 22 375
Fax + 49 30 314 21 119
Email: gruendig@inge3.bv.tu-berlin.de
Web site: www.survey.tu-berlin.de

Bernd Aschoff
Technet GmbH
Massenstrasse 14
10777 Berlin
Germany
Tel. + 49 30 215 4020
Fax + 49 30 215 4027
Email: mail@technet-gmbh.com
Web site: www.technet-gmbh.com