Generalised Procrustes Algorithms for the Conformal Updating of a Cadastral Map

Alberto Beinat and Fabio Crosilla

Summary

Referring to the current Italian cadastral map updating system, the paper describes an innovative analytical method to perform the general adjustment of the fiducial point network, and the optimal conformal insertion of new surveyed cadastral parcels into an already existing digital map. Working on the surveyors' data files containing the distance values between couples of fiducial points, the proposed procedure reciprocally fits, in the least squares sense, the various and corresponding fiducial polygons, as unitary component parts of the general fiducial network, by way of similarity transformation adjustment models computed by Generalised Procrustes algorithms. The proposed method satisfies and advantageously preserves the geometrical shape of the original surveys. This property allows the conformal mosaicking of the new surveyed cadastral entities with those ones obtained by digitisation of the original map, respecting also possible geometrical constraints of the map entities. Finally, some numerical examples show the precision and the accuracy of the method proposed.

Zusammenfassung

Mit Blick auf die derzeitigen Arbeiten zur Erneuerung des italienischen Liegenschaftskatasters wird ein innovatives Verfahren beschrieben, das sowohl die Ausgleichung des übergeordneten Netzes als auch die konforme Einpassung neu vermessener Grundstücke in bereits existierende digitale Karten erlaubt. Die in den vergangenen Jahren wiederholt durchgeführten Streckenbeobachtungen zwischen Punkten des Netzes bilden dabei Kontrollpolygone, die in optimaler Weise zusammengefügt werden müssen, so dass die Geometrie der originalen Vermessung erhalten wird. Die Methode entspricht dem Vorgehen bei der Ähnlichkeitstransformation und wird mit Hilfe des Verallgemeinerten Procrustes-Algorithmus durchgeführt.

Darüber hinaus erlaubt der Ansatz, neu vermessene Katastereinheiten konform mit solchen zu kombinieren, die aus einer Digitalisierung der Originalkarte stammen. Möglicherweise vorhandene geometrische Zwänge können gegebenenfalls berücksichtigt werden. Einige numerische Beispiele zeigen Richtigkeit und Genauigkeit der Methode auf.

1 Introduction

The earliest numerical archives of the Italian Cadastre Administration have been established, since the eighties, by digitisation of the existing cadastral maps, previously drawn on polyester films. This fact pointed out, among the others, the problem of the accuracy of the resulting parcel vertex coordinates that showed, in some circumstances, error positions of several metres.

To perform a sound update of the digital cadastral map, the Italian »Agenzia del Territorio« started at that time a virtuous process based first on the use of new topographical measurements carried out by various professional surveyors, and second, on the application of specific rules requiring, for instance, to refer the new cadastral surveys to at least three fiducial points materialised on the terrain whose original coordinate values were assigned, for the largest part of them, by the digitisation process. In this way, for the zones of recent urbanisation, a wide series of direct or indirect distance measurements between the fiducial points have been collected in the last years. By these data, the Italian Cadastre Administration, after considering a limited number of fiducial points as fixed, has locally determined the new positions of the other not fixed fiducial points, updating their corresponding coordinate values – all this according to specific significance criteria of the points taken into account and of the declared precision of the executed measurements.

However, the adopted procedure presents some disadvantages: the coordinate values of the not fixed fiducial points vary in time, and the homogenisation principle of the point coordinates might fail for a wide area, since the correction of the coordinate values acts locally and it does not rigorously propagate to the entire zone of interest. Although recent cadastral plans foresee the creation of new first and second order fiducial networks by GPS measurements, in the authors' opinion it is not proposable and extremely expensive to *ex novo* survey the entire national network of the cadastral fiducial points constituted by four millions points, and to neglect, at the same time, more than thirty million distance measurements between couples of fiducial points carried out in the last years. For this reason it becomes reasonable to propose and to experiment new analytical techniques to globally adjust all the available direct and indirect measurements acquired between the fiducial points, and constrain the corresponding geometry to a limited number of control points surveyed by GPS, in order to improve the precision of the fiducial point network and to define the coordinates of all the points into a unique common datum. The second relevant aspect is the updating problem solution of the cadastral digital map consisting in the optimal insertion of the new geometrical entities, more precise and reliable than the previously digitised ones, into the original numerical map.

The analytical procedure proposed in this paper based on the Generalised Procrustes techniques can be adopted with success, first for the rigorous adjustment of the fiducial point network, second for the optimal mosaicking of the various cadastral parcels, and third for the general and simultaneous adjustment-update of both the fiducial point network and the cadastral parcels. In the first case, by applying to each polygon (or triangle) of fiducial points surveyed in the field independent translations, rotations, and scale variations, it is possible to obtain the so-called mean shape network of the fiducial points satisfying, in the absence of gross errors, the original shape of each composing fiducial polygon. In the second hypothesis, the same analytical procedure can be properly applied for the update of the digital map by a cartographic mosaicking of the various parcels. This means, from the operational point of view, a first splitting of all the digital parcels of the original cadastral map and their subsequent recomposition with the insertion of the new ones surveyed in the field with better precision and reliability. Also in this case, the Generalised Procrustes algorithm allows the analytical least squares adaptation of the whole set of the parcels preserving their original shape and satisfying some geometrical constraints of the various entities. Finally, given the strict correlation and the geometrical interdependence existing between the fiducial points and the vertexes of the surveyed parcels, the proposed procedure of cartographic updating can be simultaneously applied to the rigorous adjustment of the fiducial point networks and to the optimal integration of the surveyed parcels with those digitised from the original cadastral maps. This last procedure of *general updating* is, from the geometrical point of view, the most advanced and optimal one, since it benefits of the constraints and of the high correlation existing between the geometrical position of the fiducial points and the cartographic location of the various parcels.

2 An Original Proposal for the Rigorous Conformal Adjustment of the Fiducial Polygons

Every time a parcel is surveyed in order to update or verify the digital archives, the Italian Cadastre rules prescribe that the corresponding field measurements must be recorded, processed, and uploaded into the numerical archives by way of a specific freely available computer procedure, known as the »Pregeo« program. As outlined before, the current specifications also state that all the point coordinates describing the new shape and position of the parcels must be measured with respect to a local reference frame materialised by at least three fiducial points (fiducial triangle) or more (fiducial polygon) whose perimeter completely or almost completely contains the geometric entity to survey (Fig. 1). From tasks

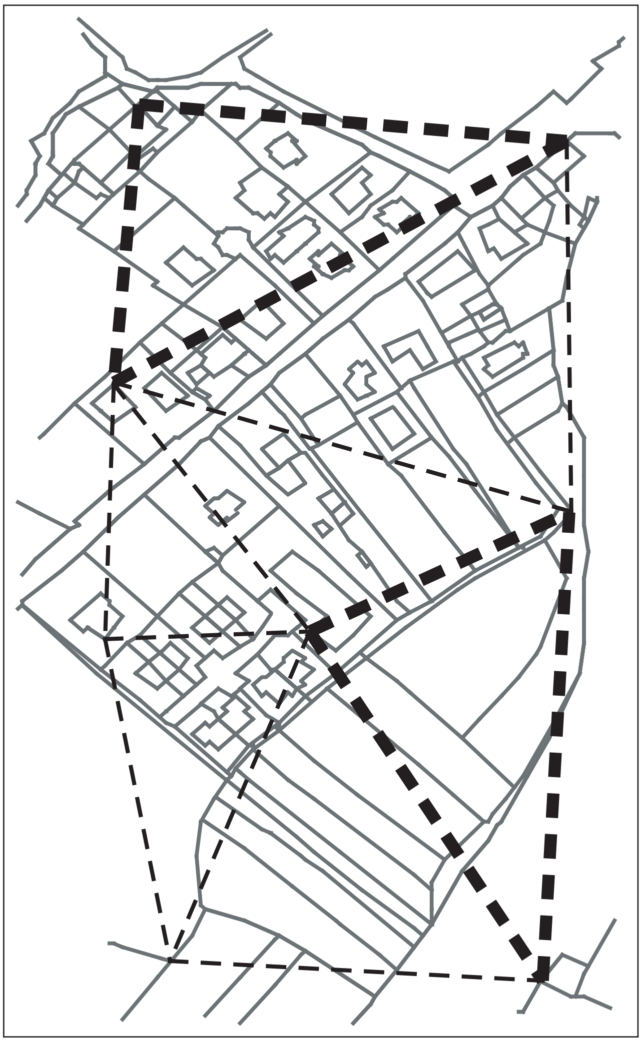

Fig. 1: Fiducial point network superimposed to the cadastral map. Two fiducial triangles are evidenced

of this kind it is then possible to extract the reciprocal distance values existing between pairs of fiducial points, as determined by the actual field measurements, and to log them into a file. Univocal and rigorous naming standards are used to define the fiducial point names at a national level; this simplifies and makes straightforward every topological connection required by automated processing. All this specific information available in digital form represents a precious opportunity to attempt the solution of the first main cadastral map problem listed above, that is the rigorous re-adjustment of the fiducial point network.

The fundamental philosophy of the proposed approach is that it aims at preserving the original shape of every fiducial polygon (hence the term »conformal«). Its geometric configuration, in fact, has been determined by a variable set of angular and distance measurements performed among the various fiducial points and the parcel vertexes and results from the consequent internal measurements adjustment. All this furnishes an »invisible skeleton« to the whole fiducial polygon, giving the reason of its size and shape. The procedure that performs these transformations is based on the Generalised Procrustes Analysis, already employed to solve a classical photogrammetric problem (Crosilla and Beinat 2002). Other recent geodetic applications of the Procrustes analysis are due to Grafarend and Awange (2000). Procrustes transformations are widely applied techniques for the shape analysis, the multifactorial analysis, and for the multidimensional scaling problems (e. g. Borg and Groenen 1997, Dryden and Mardia 1998). They make it possible the least squares fit of two or more data matrices by searching a proper set of rotation matrices, scaling factors and translation vectors in order to satisfy a predefined objective function. The procedure consents to solve complex multidimensional problems by way of a rigorous sequence of simple elementary bi-dimensional or tri-dimensional transformation solutions. One of the major advantages of the proposed method is that it does not require any linearization of the equation system involved. As a direct consequence, the need for approximated values of the unknown parameters to be estimated is never required. The full process has been structured in two moments. In the first step, the goal is to establish the most probable shape for a given fiducial triangle (or polygon), taking into account the variability of the different surveys of it carried out during the time. As a result of the data analysis performed in this stage it is also possible to compute an index of accuracy for every fiducial polygon that can be used as a weight in the subsequent general adjustment. To this aim, fiducial polygons of the network that were never determined directly in the field or measured with poorly accurate methods or instruments assume a proportionally reduced weight. The second step concerns the general and simultaneous adjustment of the representative shapes *(consensus)* of all the fiducial polygons. By the same algorithms previously used to define the most probable shape of every set of homologous fiducial polygons, the various representative polygons are linked one to the others by their common points in order to rigorously recompose the network of the fiducial points. The procedure individually rotates, translates, and residually scales all the polygons in order to obtain their reciprocal best fit condition respecting the constraints imposed by an existing control network (e. g. GPS). This state is achieved when the sum of the squared residual distances between the homologous fiducial points belonging to the various polygon determinations and their mean position satisfies a minimum condition. Finally, the definitive position of one specific fiducial point is obtained by averaging the coordinate definitions of the same point after having rotated, translated, and residually scaled in optimal way the different fiducial polygons that contain it. During the adjustment process the vertexes of the first and second order of the fundamental cadastral network and the vertexes of the National Geodetic network act as fixed points. In this way not only the adjusted network is constrained to the control points letting only local adaptations among the fiducial polygons, but it is also possible to perform a local datum exchange from one reference frame to another by a similarity transformation restricted to the typical extent of a fiducial polygon (of the order of hundreds of metres).

In the actual implemented procedure, both the steps of the adjustment operate at the same time. It results in fact the same process, with the same rules, algorithms, and goals. It has been divided, in the explanation, in order to make it more comprehensible.

3 The mathematical model

We focus first on the problem of establishing the most probable shape of a given fiducial triangle (or polygon) resulting from a series of geometric determinations of it, considering their different accuracy and their datum variability. Let $\mathbf{A}_1 \ldots \mathbf{A}_m$ be a set of m data matrices containing each one, by rows, the point coordinates of the same set of p fiducial points belonging to a given fiducial polygon $\mathbf{A}$, as determined by m different field surveys carried out at different times. For more generality we can assume tri-dimensional point coordinates (two planimetric components x and y, and the height z), consequently every $\mathbf{A}_i$ has dimension $p \times 3$ (Fig. 2). At present, of course, this does not represent the situation of Italy, but the course to develop a 3-D cadastral map in the future is going to be undertaken. The further assumption is that the coordinates of every fiducial polygon $\mathbf{A}_i$ $(i = 1 \ldots m)$, are defined in a local arbitrary datum that can be related to the others and to the general map reference frame by an appropriate similarity transformation.

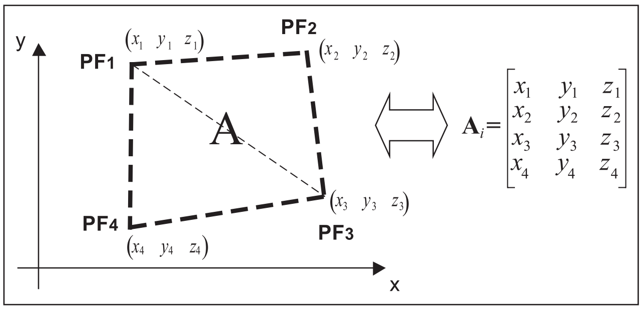

Fig. 2: A generic fiducial polygon and the corresponding coordinate matrix

The most probable shape of mean size $\hat{\mathbf{A}}$ of a given fiducial polygon $\mathbf{A}$ can not be directly computed by simply averaging the corresponding point coordinate values, since there could be systematic components among the different fiducial polygon determinations $\mathbf{A}_i$ $(i = 1 \ldots m)$ (different origin, orientation, and scale) that must be removed at first. Therefore, a proper unknown similarity transformation $\{c, \mathbf{T}, \mathbf{t}\}_i$ – where $\mathbf{T}$ is an orthogonal 3-d rotation matrix, $\mathbf{t}$ a translation vector, and c a scaling factor – must be computed for every $\mathbf{A}_i$ $(i = 1 \ldots m)$ in order to contemporaneously fit it to all the others $\mathbf{A}_k$ $(k = 1 \ldots m; k \neq i)$. This leads, unless an individual residual error matrix $\mathbf{E}_i$ ($j=1\ldots m$), to the unknown real configuration matrix $\mathbf{A}$

$$\begin{aligned}\mathbf{A} &= c_1\mathbf{A}_1\mathbf{T}_1 + \mathbf{j}\mathbf{t}_1 + \mathbf{E}_1 \\ &= c_i\mathbf{A}_i\mathbf{T}_i + \mathbf{j}\mathbf{t}_i + \mathbf{E}_i \\ &= c_m\mathbf{A}_m\mathbf{T}_m + \mathbf{j}\mathbf{t}_m + \mathbf{E}_m\end{aligned} \quad . \tag{1}$$

In Equation (1) $\mathbf{j}$ is an auxiliary column vector of size $p \times 1$ with all the components set to 1, $\text{vec}(\mathbf{E}_i)$ is normally distributed with zero mean and covariance matrix $\Sigma = \sigma^2\mathbf{I}$.

Among the infinite similarity transformation solutions satisfying Equation (1) – excluding the trivial ones for which $c_i = 0\ \forall i$ – we look for that one which fulfils the following least-squares condition:

$$S = \text{tr}\sum_{i<k}^{m}\left[\left(c_i\mathbf{A}_i\mathbf{T}_i + \mathbf{j}\mathbf{t}_i\right) - \left(c_k\mathbf{A}_k\mathbf{T}_k + \mathbf{j}\mathbf{t}_k\right)\right]^T \left[\left(c_i\mathbf{A}_i\mathbf{T}_i + \mathbf{j}\mathbf{t}_i\right) - \left(c_k\mathbf{A}_k\mathbf{T}_k + \mathbf{j}\mathbf{t}_k\right)\right] = \min \quad . \tag{2}$$

It represents the simultaneous best fit among $\mathbf{A}_1\ldots\mathbf{A}_m$, that is the result of the well-known Generalised Procrustes Analysis problem (GPA) described and solved by Gower (1975), Ten Berge (1977) and Goodall (1991). After the minimum condition of Formula (2) is numerically satisfied, the estimate $\hat{\mathbf{A}}$ of the consensus configuration $\mathbf{A}$ is given by:

$$\hat{\mathbf{A}} = \frac{1}{m}\sum_{i=1}^{m} c_i\mathbf{A}_i\mathbf{T}_i + \mathbf{j}\mathbf{t}_i \ . \tag{3}$$

We refer to the above mentioned literature for the theoretical aspects and to Appendix B for the description of a problem solution method. Fig. 3 aims to provide a graphical interpretation of the GPA. Four geometric configurations ($\mathbf{A}_i$, $i=1\ldots4$) of the same fiducial polygon $\mathbf{A}$, produced by four separate surveys are represented in Fig. 3-a with their differences deliberately exaggerated. In Fig. 3-b the various polygons are centred one over the others by way of a proper translation, and in Fig. 3-c they are rotated and scaled up to their maximal agreement. Fig. 3-d illustrates the final shape assumed by each of them, and their relationship with their most probable shape of mean size $\hat{\mathbf{A}}$, computed averaging the corresponding point coordinates of the transformed original matrices $\mathbf{A}_i$ ($i=1,2,3,4$).

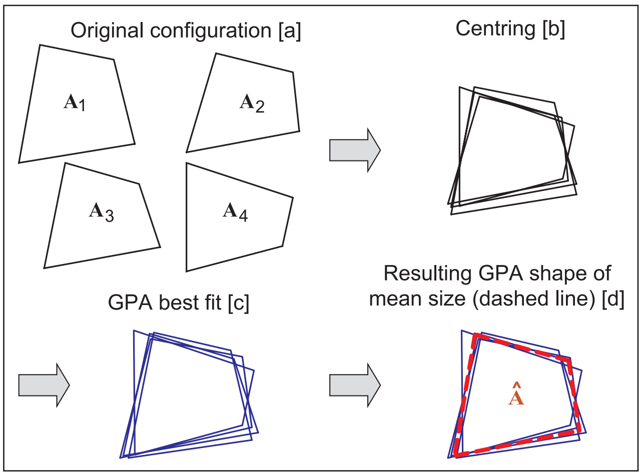

Fig. 3: Geometric representation of the GPA process

Once all the most probable shapes of mean size of every polygon have been determined, taking into account their relative accuracy, the following problem requires to fit each polygon to its adjacent ones, by way of the common fiducial points. This task presents an approach similar to the photogrammetric block adjustment by independent models for which a Procrustes solution has been recently developed (Crosilla and Beinat 2002). The mathematical formulation of the general adjustment begins with the GPA model extended to consider partially weighted configurations. The general form of the Procrustes problem assumes therefore the following expression (Commander 1991):

$$S = \sum_{i<k}^{m}\text{tr}\left(c_i\mathbf{A}_i\mathbf{T}_i + \mathbf{j}\mathbf{t}_i - c_k\mathbf{A}_k\mathbf{T}_k - \mathbf{j}\mathbf{t}_k\right)^T \mathbf{D}_i\mathbf{D}_k\left(\sum_{j=1}^{m}\mathbf{D}_j\right)^{-1} \left(c_i\mathbf{A}_i\mathbf{T}_i + \mathbf{j}\mathbf{t}_i - c_k\mathbf{A}_k\mathbf{T}_k - \mathbf{j}\mathbf{t}_k\right) = \min \quad , \tag{4}$$

where $\mathbf{D}_i$ and $\mathbf{D}_k$ are general diagonal weighting matrices of size $p \times p$. It can be demonstrated (e. g. Borg and Groenen 1997) that the condition expressed by Formula (4) is equivalent to:

$$S = \sum_{i=1}^{m}\text{tr}\left(c_i\mathbf{A}_i\mathbf{T}_i + \mathbf{j}\mathbf{t}_i - \hat{\mathbf{A}}\right)^T \mathbf{D}_i\left(c_i\mathbf{A}_i\mathbf{T}_i + \mathbf{j}\mathbf{t}_i - \hat{\mathbf{A}}\right) = \min \ , \tag{5}$$

where $\hat{\mathbf{A}}$, called the geometric centroid, is given by:

$$\hat{\mathbf{A}} = \left(\sum_{k=1}^{m}\mathbf{D}_k\right)^{-1}\sum_{i=1}^{m}\mathbf{D}_i\left(c\mathbf{A}_i\mathbf{T}_i + \mathbf{j}\mathbf{t}_i\right) \ . \tag{6}$$

$\hat{\mathbf{A}}$ represents the least squares estimation of the real configuration matrix $\mathbf{A}$ (Crosilla and Beinat 2002). This second formulation can be managed and computed more easily than the previous expressed by Formula (4). Equations (5) and (6), in fact, are computed iteratively, and matrices $\mathbf{A}_i$ are continuously updated, until a predefined convergence threshold is reached. This event represents the GPA best fit solution. The algorithm that makes possible to update, at each iteration, every data matrix $\mathbf{A}_i$ ($i=1\ldots m$) with respect to matrix $\hat{\mathbf{A}}$, is reported in Appendix A.

One fundamental advantage of the generalised formulation for the weighted case is the capability to account for situations of missing corresponding points between matrices. To this aim, an efficient solution is due to Commander (1991). Every $\mathbf{D}_i$ can be considered as the product of a proper weight matrix $\mathbf{P}_i$ by a Boolean diagonal matrix $\mathbf{M}_i$:

$$\mathbf{D}_i = \mathbf{M}_i\mathbf{P}_i = \mathbf{P}_i\mathbf{M}_i \ . \tag{7}$$

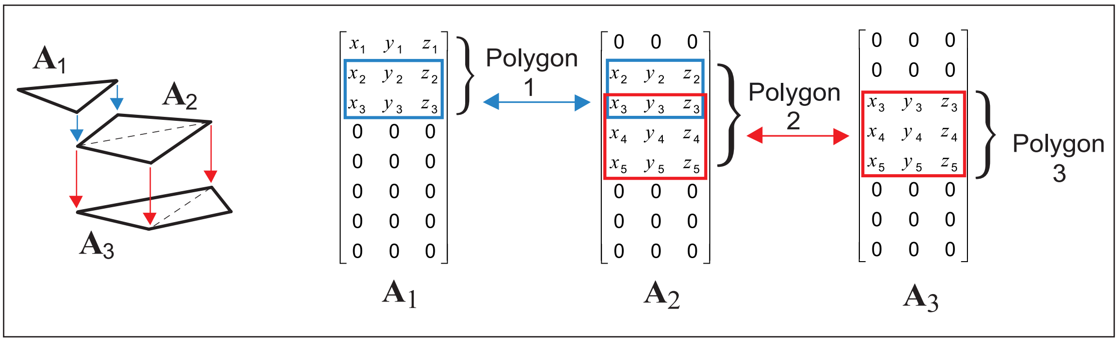

Fig. 4: Fiducial polygons, matrix description, and fiducial point correspondences

$\mathbf{M}_i$ is automatically defined and associated to every matrix $\mathbf{A}_i$. Its diagonal components are 1 where the corresponding elements (rows) of $\mathbf{A}_i$ are effectively defined, and zero in all the other cases. Referring to Fig. 4, the diagonals of the $\mathbf{M}_i$ matrices associated to the corresponding $\mathbf{A}_i$ are:

$$\mathrm{diag}(\mathbf{M}_1) = [1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0],$$
$$\mathrm{diag}(\mathbf{M}_2) = [0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0],$$
$$\mathrm{diag}(\mathbf{M}_3) = [0 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0 \quad 0].$$

To better understand how the GPA problem with missing points can be set out, we have represented all the matrices $\mathbf{A}_i$ with the same size, that is the same coordinate dimension and an equal number of points, although not all are actually defined. Ideally, every $\mathbf{A}_i$ contains as many rows as the total number of fiducial points to be adjusted. For the algorithm implementation this assumption does not represent a drawback. In fact, the apparent waste of computer memory resources can be avoided introducing and managing variable size arrays by way of usual programming techniques. The software implementation deserve some additional remarks. Practically, the two stages in which the network adjustment has been divided, that is the initial identification of the most probable size and shape of every fiducial polygon configuration, and the following reciprocal best fit of the different ones, are performed simultaneously. Every measured polygon assumes a proper global weight as a function of the accuracy and of the survey techniques adopted for its determination.

Regarding the geometrical constraints, they are represented as fixed point coordinates collected into a further $\mathbf{A}_{m+1}$ matrix whose values are kept unchanged during the computations. Borderline alignments are performed by orthogonal projection of the border points towards the corresponding nearest border segments. The resultant constraint matrix $\mathbf{A}_{m+2}$ contains the intersection point coordinates defined in this way that are updated after every iteration of the GPA process. The function in Equation (4) is therefore extended to account for the constraints matrices introduced.

4 Case Studies

4.1 Application of the proposed method to a simulated example

A rigorous evaluation of the Procrustes adjustment model capabilities to solve problems relative to the cadastral map recomposition in the presence of different kinds of errors can be done if the true measurement values are available. For this purpose, a real situation has been artificially created considering a network of fiducial points with fixed coordinate values. Afterwards, various kinds of errors have been inserted into the original data simulating in this way incorrect measurements surveyed in the field.

4.2 Description of the simulated network

A two dimensional network of 30 vertexes has been considered (Fig. 5). In order to simulate a real spatial distribution of the fiducial points, the vertexes have been located with a reciprocal distance varying between 227 m and 1153 m; the total area covered by the network corresponds to 5.7 Km². The network is composed of a series of triangles and polygons whose vertexes represent in reality the fiducial points used by the technicians to reference the various surveys. In this way 70 polygons have been artificially generated and some of them are characterised by common fiducial points. Furthermore, some fixed points have been identified within the network, that

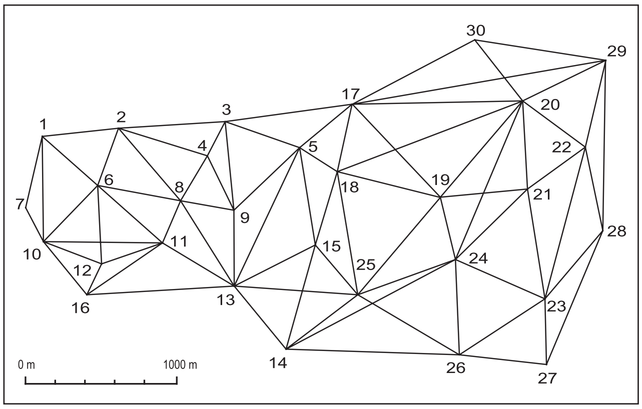

Fig. 5: The simulated fiducial point network

is points of high importance for the Cadastre Administration, or points whose coordinates have been determined with a high precision (for instance by GPS). Their choice depends on:

- the spatial distribution of the fixed points that must be homogeneous over the entire network;
- the spatial density of the fixed points that must satisfy the technical specifications of the Cadastre for a second order network;
- the number of polygons having in common the same fixed points.

According to these conditions, three groups of fixed points have been identified:

Group 1: constituted by 9 fixed points, that is 1.59 points per Km^2, partially distributed along the perimeter of the network and partially in the internal part;

Group 2: constituted by 7 fixed points, that is 1.19 points per Km^2, distributed along the perimeter of the network;

Group 3: constituted by 4 points, that is 0.70 points per Km^2, distributed along the perimeter of the network.

4.3 Error attribution for the simulated network

Before the adjustment, particular error models have been applied to the original file containing the true coordinates of the various polygons: in this way a new set of data files simulating the effective measurements has been obtained. According to the kind of errors applied, five different classes and subclasses of analysis have been generated:

1) measurements affected only by systematic errors (S);
2) measurements contemporarily affected by systematic and random errors (S and R) which can be partitioned into three different subclasses:
   2a) prevalence of the systematic error component with respect to the random one (S > R);
   2b) equivalence between the systematic and the random error components (S = R);
   2c) prevalence of the random error component with respect to the systematic one (R > S);
3) measurements affected only by random error components (R).

Systematic errors have been generated by randomly and rigidly scaling the single fiducial polygons: the greatest scale perturbations are within the range of ± 100 ppm. Properly speaking, the systematic effects can be only evidenced at the local level. For each class and subclass of analysis, three different error simulations have been executed. The total number of datasets is therefore equal to 15; in this way a significant number of possible situations has been originated. Every dataset has been then adjusted applying the Procrustes method considering for every computation the scheme of fixed points characterising one of the three mentioned groups. In this way, 45 different Procrustes adjustments have been performed and the results compared with the original true values (Clerici 2002).

4.4 Numerical results

The numerical experiments have been executed with the aim to verify the capability of the Procrustes algorithm to compensate for the various kinds of error. The following Tab. 1 reports the mean values, the medians, and the root-mean square errors for a specific error simulation, referring to each of the fixed point groups. Analogous results have been reached for the other two cases of error generation. An analysis of the results contained in Tab. 1, in particular for what is referred to Group 1, allows to conclude that for the Procrustes analysis the mean and the median values for the x and y coordinate components never differ for values greater than 0.7 cm in absolute value, indicating in this way a good symmetry of the

Tab. 1: Summary of the mean, median, and rmse values of the coordinate differences for the Procrustes adjustment method [cm].

Error type			Mean			Median			RMSE		
			Group 1	Group 2	Group 3	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
S		x	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.1	0.0
		y	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
S + R	S > R	x	0.2	−0.3	0.6	0.0	−0.4	0.2	1.5	1.7	3.3
		y	−0.3	−0.5	−1.6	−0.3	−0.6	−1.0	1.2	1.8	4.3
	S ≈ R	x	1.2	−0.9	−0.3	0.5	−1.4	−0.5	4.3	4.5	8.9
		y	−0.8	1.3	−3.3	−0.2	0.7	−2.1	4.1	4.4	6.3
	S < R	x	1.0	−4.7	−12.7	−0.5	−2.0	14.7	15.3	17.0	17.5
		y	0.6	−6.8	−5.9	0.3	−6.5	−3.9	12.1	15.1	22.2
R		x	1.1	−0.7	0.9	0.5	−1.7	0.9	4.5	4.7	8.6
		y	−0.7	−0.4	−6.0	−0.1	−0.2	−5.5	3.9	4.2	8.6

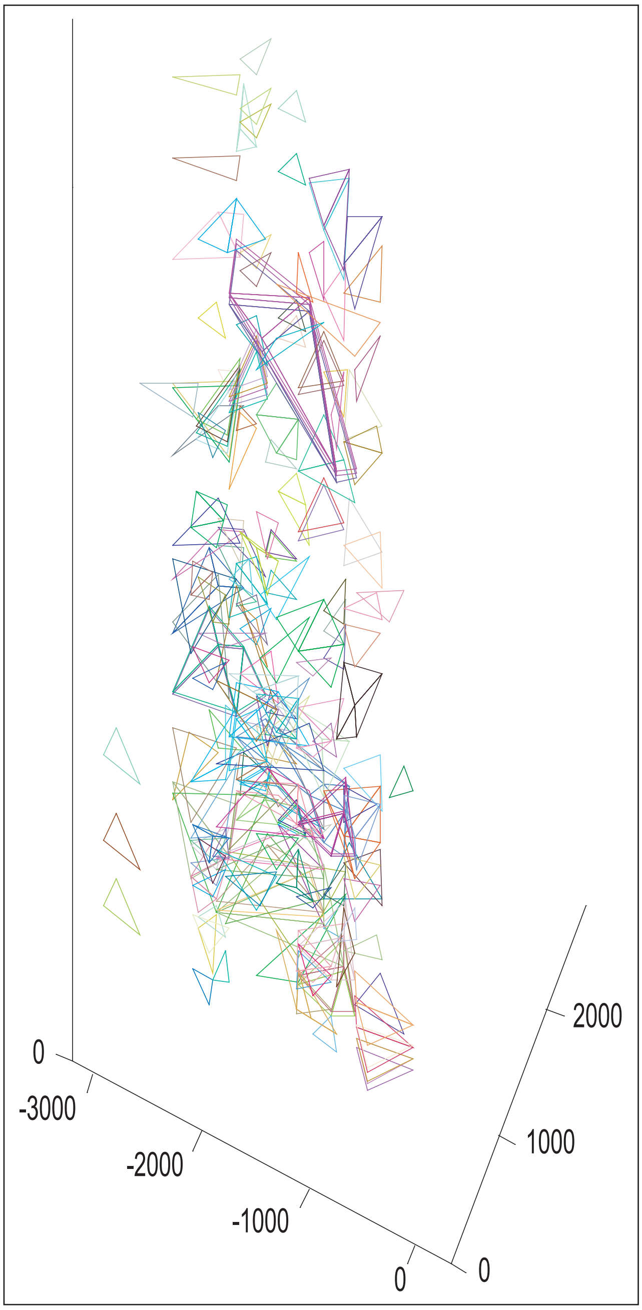

Fig. 6: The 205 adjacent, partially or totally overlapping fiducial polygons composing the network.

residual distribution. Furthermore, it is possible to observe that in the case of the only presence of systematic errors, Procrustes adjustment leads to mean and median values equal to zero for all the three groups of fixed points. In this case, the residuals follow a perfectly symmetrical distribution without any systematic error component. If only random errors are present, the values of the differences between means and medians remain comparable; this is true also for the values of the root mean square error, but is limited to the families of fixed points 1 and 2 that consider a sufficient number of fixed points for the whole network.

As the previous results show, the Procrustes technique seems very powerful in the presence of systematic errors and can be used also in the case in which only random errors are present; this is true if and only if the fixed points adequately cover the network taken into account.

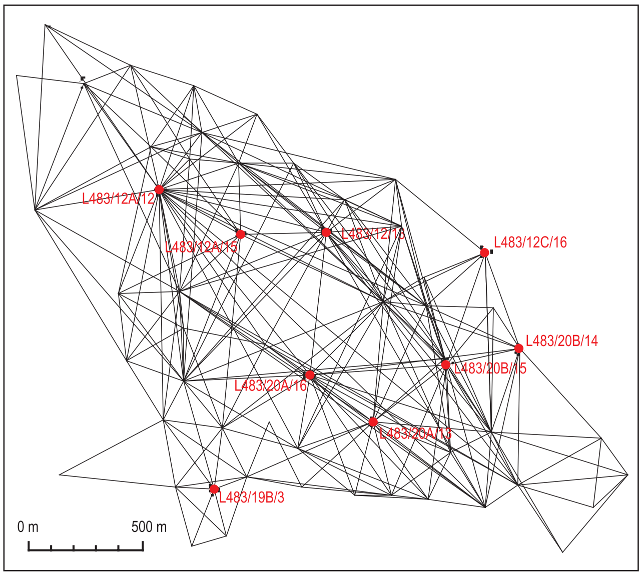

Fig. 7: The fiducial point network used for the test, with the reciprocal distances measured between couples of fiducial points. The 9 check points are evidenced.

4.5 A numerical experiment to a real example

The Udine branch of the Italian »Agenzia del Territorio« made available the coordinates values and the description of the fiducial points surveyed since 1990 in the North-West part of the Udine municipality. The network is constituted by 64 fiducial points covering an area of almost 2.8 Km^2; the total number of polygons used for the various surveys is equal to 205 (Fig. 6 and 7).

The initial experimental phase consisted in the GPS measurement of 18 fiducial point positions. Of these points, 9 were considered like fixed points and homogeneously distributed over the entire network, while the remaining 9 have been used to control the results. The data have been then adjusted using the Procrustes algorithm. The precision of the adjustment with the Procrustes method is referred to the values of the root mean square errors computed from the coordinates of the adjusted fiducial points and their mean value. A mean value of precision equal to ± 0.051 m has been obtained showing in this way the mutual adaptation capability of the various polygons permitted by the application of a different scale corrective factor. The accuracy of the results is evaluated analysing the differences between the »true« coordinates of the check points measured in the field by GPS and the coordinates of the same points obtained by the adjustment (Tab. 2)

These first results confirm the capability of the Procrustes method to perform the general adjustment of the fiducial point polygon network with an accuracy that, in module, is really promising for the cadastral expectations.

Tab. 2: Residuals between the GPS coordinates and the adjusted ones, relative to the check points [cm]

Check point id	North	East	Module
L483/12/13	–8.9	9.3	12.9
L483/12A/12	–13.6	27.9	31.0
L483/12A/15	2.8	26.7	26.8
L483/12C/16	2.6	–4.1	4.9
L483/19B/3	–2.7	–8.5	8.9
L483/20A/13	–7.7	–6.0	9.8
L483/20A/16	–10.7	–14.9	18.3
L483/20B/14	–3.7	2.7	4.6
L483/20B/15	–7.3	2.9	7.9

5 Concluding Remarks

The paper describes an original analytical method to perform the updating of a cadastral digital map using new measurements executed in the field by professionals with enough precision and reliability. The proposed method, based on a series of algorithms coming from the Procrustes analysis, makes it possible to join and to fit, in an optimal and rigorous way, the various fiducial polygons surveyed in the field without modifying their original shape. The same procedure can be extended to the conformal mosaicking of new surveyed cadastral entities with those obtained by the digitisation of the original map, satisfying also possible geometrical constraints of the various elements.

The experimental results show the good performances of the method proposed, particularly when residual systematic errors in the field measurements are present. The Procrustes are able to compensate all the systematic components and behave as a traditional adjustment procedure in the presence of only random errors.

Future research will be devoted to the reliability aspects of the Procrustes method in order to study the possibility to detect wrong surveys by the application of proper statistical tests of shape analysis.

Acknowledgments

The authors are grateful to Dr.-Ing. Friedhelm Krumm for the German translation of the summary. Special thanks to the Italian »Agenzia del Territorio« for the collaboration and the data kindly provided.

References

Beinat, A., Crosilla, F.: Generalised Procrustes Analysis for Size and Shape 3-D Object Reconstructions. Optical 3-D Measurement Techniques, Wien 1–4 October, V, 345–353, 2001.

Borg, I., Groenen, P.: Modern multidimensional scaling: theory and applications. Springer-Verlag, New York, 337–379, 1997.

Commandeur, J.J.F: Matching configurations. DSWO Press, Leiden University, III, M&T series, 19, 13–61, 1991.

Clerici, A.: Una proposta innovativa per la ricomposizione cartografica del catasto basata sulle tecniche di analisi procustiana. Degree thesis, Faculty of engineering, University of Udine, 2002.

Crosilla, F., Beinat, A.: Use of Generalised Procrustes analysis for photogrammetric block adjustment by independent models. ISPRS Journal of Photogrammetry & Remote Sensing, Elsevier, 56 (3), 195–209, 2002.

Dryden, I.L., Mardia, K.W.: Statistical shape analysis. John Wiley & Sons, Chichester, England, 83–107, 1998.

Goodall, C.: Procrustes methods in the statistical analysis of shape. Journal Royal Stat. Soc., Series B-Methodological, 53 (2), 285–339, 1991.

Gower, J.C.: Generalized Procrustes analysis. Psychometrika 40 (1), 33–51, 1975.

Grafarend, E.W., Awange, J.L.: Determinations of vertical deflections by GPS/LPS measurements. ZfV 8, 279–288, 2000.

Schoenemann, P.H., Carroll, R.: Fitting one matrix to another under choice of a central dilation and a rigid motion. Psychometrika, 35 (2), 245–255, 1970.

Ten Berge, J.M.F.: Orthogonal Procrustes rotation for two or more matrices. Psychometrika, 42 (2), 267–276, 1977.

Authors' address

Dr. Ing. Alberto Beinat
Prof. Fabio Crosilla
Department of Georesources and Territory
University of Udine
Via Cotonificio, 114
I-33100 Udine (Italy)
Tel.: ++39 0432 558702
Fax: ++39 0432 558700
beinat@dgt.uniud.it
crosilla@dgt.uniud.it

Appendix A

The similarity transformation parameter estimation between two matrices **A** and **B** with the same dimension, indefinitely translated, rotated, and scaled with respect to each other, can be directly solved with the following method (Schoenemann and Carroll 1970). Given an (optional) diagonal weighting matrix **P**, containing global weighting factors associated to every point of **A** and **B**, and **j** a predefined auxiliary unitary vector ($\mathbf{j} = [1_1 \ldots 1_p]$), where p is the number of points of **A** and **B**, by substituting $\mathbf{A}_W = \mathbf{P}^{½}\mathbf{A}$, $\mathbf{B}_W = \mathbf{P}^{½}\mathbf{B}$, $\mathbf{j}_W = \mathbf{P}^{½}\mathbf{j}$ and computing the singular value decomposition of the following matrix product:

$$\mathbf{A}_W^T \left(\mathbf{I} - \frac{\mathbf{j}_W \mathbf{j}_W^T}{\mathbf{j}_W^T \mathbf{j}_W} \right) \mathbf{B}_W = \mathbf{V}\mathbf{D}_s\mathbf{W}^T \ , \tag{8}$$

under the orthogonality condition $\mathbf{T}^T\mathbf{T} = \mathbf{I}$, the parameters **t**, **T** and c of the similarity transformation that converts **A** over **B** are obtained by:

1. the rotation matrix **T**:

$$\mathbf{T} = \mathbf{V}\mathbf{W}^T ; \tag{9}$$

2. the global scale factor c:

$$c = \mathrm{tr}\left[\mathbf{T}^T \mathbf{A}_W^T \left(\mathbf{I} - \frac{\mathbf{j}_W \mathbf{j}_W^T}{\mathbf{j}_W^T \mathbf{j}_W} \right) \mathbf{B}_W \right] \Big/ \mathrm{tr}\left[\mathbf{A}_W^T \left(\mathbf{I} - \frac{\mathbf{j}_W \mathbf{j}_W^T}{\mathbf{j}_W^T \mathbf{j}_W} \right) \mathbf{B}_W \right] ; \tag{10}$$

3. the translation vector **t**:

$$\mathbf{t} = \left(\mathbf{B}_W - c\mathbf{A}_W \mathbf{T} \right)^T \frac{\mathbf{j}_W}{\mathbf{j}_W^T \mathbf{j}_W} \ . \tag{11}$$

Appendix B

The solution of the GPA problem can be achieved by an iterative procedure proposed by Gower (1975) and improved by Ten Berge (1977). It operates the sequential updating of the various matrices $\mathbf{A}_i$ up to their maximum agreement, splitting the problem in three main phases:

a) Translation problem: solved by means of an initial column centring of matrices $\mathbf{A}_i$, i. e. each column is reduced to the centre of gravity of its values.

b) Rotation problem: solved by the following iterative procedure:

Step 1: $\mathbf{A}_1$ is rotated with respect to $\mathbf{B} = \sum_{j=2}^{m} \mathbf{A}_j$, so to obtain $\mathbf{A}_1^{(1)} = \mathbf{A}_1 \mathbf{T}_1^{(1)}$;

Step 2: $\mathbf{A}_2$ is rotated with respect to $\mathbf{B} = \mathbf{A}_1^{(1)} + \sum_{j=3}^{m} \mathbf{A}_j$, so to obtain $\mathbf{A}_2^{(1)} = \mathbf{A}_2 \mathbf{T}_2^{(1)}$;

Step m: $\mathbf{A}_m$ is rotated with respect to $\mathbf{B} = \sum_{j=1}^{m-1} \mathbf{A}_j^{(1)}$, so to obtain $\mathbf{A}_m^{(1)} = \mathbf{A}_m \mathbf{T}_m^{(1)}$;

Step m + 1: $\mathbf{A}_1^{(1)}$ is rotated in turn with respect to $\mathbf{B} = \sum_{j=2}^{m} \mathbf{A}_j^{(1)}$, so to obtain $\mathbf{A}_1^{(2)} = \mathbf{A}_1^{(1)} \mathbf{T}_1^{(2)}$;

and so on, up to a convergence for the estimation of rotation, whereby $^{(i)}$ above represents the iteration number, and **B** is a temporary matrix recomputed at every step. The specific unknown rotation matrix $\mathbf{T}_i$ that fits the generic matrix $\mathbf{A}_i$ over the temporary matrix **B**, is reported in Appendix A.

(c) Scale factor problem: solved by the computation of the correlation matrix of vec($\mathbf{A}_i$), where the vec operator returns the vector format of the matrix $\mathbf{A}_i$ by columns. Assuming the updated values for the coordinate matrices $\mathbf{A}_i$ ($j = 1 \ldots$m), as they result from the rotation computations, the scale factor c_i, corresponding to the generic matrix $\mathbf{A}_i$, is given by:

$$c_i = \left\{ \left(\sum_{j=1}^{m} \left\| \mathbf{A}_j \right\|^2 \right) \Big/ \left\| \mathbf{A}_i \right\|^2 \right\}^{\frac{1}{2}} f_i \ ,$$

where f_i is the *i*-th element of the eigenvector **f**, corresponding to the greatest eigenvalue of the correlation matrix of vec($\mathbf{A}_i$).
The computation sequence of rotation and scaling is iterated up to a convergence of both parameters (general convergence). At this stage the estimate of the consensus configuration is computed averaging the final component values of the $\mathbf{A}_i$ matrices, as they result at the conclusion of the iterative updating.

$$\hat{\mathbf{A}} = \sum_{j=1}^{m} \mathbf{A}_j^{(\mathrm{final})} \ .$$