AUTOMATIC POINT MATCHING OF GIS GEOMETRIC FIGURES

A. Beinat, F. Crosilla, E. Sossai

Dept. of Georesources and Territory, University of Udine, Via Cotonificio 114, I-33100 Udine (Italy)
Tel. ++39 0432 558702 Fax ++39 0432 558700 - (beinat, crosilla, sossai)@dgt.uniud.it

KEY WORDS: Cartography, GIS, CAD, Automation, Identification, Databases

ABSTRACT:

GIS and digital mapping operations frequently require the automatic comparison and superimposition of geometric figures represented by sets of vertex coordinates supported by structural and topological information. When the configurations are not structured, that is the only vertex coordinates of the figures are available, manual intervention is needed in order to establish correspondences among the different geometries.

To overcome this limitation, an automatic method has been developed to detect the correspondences between two or more equivalent sets of unlabeled points, representing n-dimensional geometric figures. The proposed technique performs a geometrical analysis of the adjacency matrices of the point configurations, in order to identify, for each one, the vertex of maximal asymmetry. A pairwise comparison of the sorted components of the adjacency matrix relative to these vertices, leads to the identification of the point correspondences. A directly-computed Procrustes conformal transformation is then applied to the geometric figures in order to achieve their optimal alignment.

Also in case of geometric entities included into another, the problem solution starts trying to find some minimal asymmetric sub-configurations (kernels) that are similar in both figures. A Procrustes superimposition of these corresponding kernels is then applied, and extended to the remaining points of the included configuration. A shape test is finally executed in order to identify the best solution. Specific geometric rules and filters are implemented to optimise the computation process.

The method has been successfully tested on cadastral cartographic matching problems. In addition, it is suitable for a wider range of possible applications, like CAD/CAM, computer vision and reverse engineering.

1. INTRODUCTION

Most of the GIS tools used to handle a digital map, and many CAD/CAM applications, are not currently provided of specific functions able to automatically identify the geometrical correspondence between a specific geometrical figure, and that part of a more general drawing containing it. This is not true just for the cases in which the reference data files are appropriately structured.

Problems of this kind often arise in digital mapping and in GIS, where, given a cartographic element, for instance a cadastral parcel, the problem is to automatically recognize it within a general map. An analogous case is given by the automatic research of a predefined structural or mechanical component unit within some CAD/CAM files containing the entire structure or mechanism.

In particular cases, the solution of these problems is possible through the a priori definition of some topological relationships among the object points, or from the availability of structural information, explanatory of how the points are mutually connected in order to represent specific shapes. It is established, for instance, that a collection of points, properly ordered and joined according to a certain rule, define a cadastral parcel, and the comparison between the specific structured dataset and the cadastral map is carried out. In this case, the map is considered like a cluster of unit parcels defined by a proper series of points organized in a topologically compatible manner.

To structure the data archives in such a way, it requires a heavy and time spending manipulation, however not sufficient to solve the problems of identification and automatic comparison in the various ways in which they occur. It is impossible, for instance, to compare a new object with the existing archive, if the topology of the new object is unknown, undefined, or incompatible.

Therefore, when such a kind of information is not available, or is incompatible, it is only possible to assume the fundamental geometrical information, i.e. the coordinates of the composing points that graphically define the involved geometrical configurations.

Problems of this type have already attracted the attention of researchers mainly working in the field of the computer graphics. One fundamental contribution to the problem solution is due to Ullman (1979), who recognised the importance of using the stiffness of the existing constraints between two vertices, with the aim of identifying the correspondence among point sets. Later, many authors, being inspired by Ullman's work, developed algorithms, using the weighed proximity matrix as a function of the computed distance lengths among the points. In particular, Scott and Longuett-Higgins (1991) have determined the correspondence by the spectral decomposition of the proximity matrix relative to point configurations under study. Shapiro and Brady (1992) reached the objective through the comparison of the modal structures deduced from each proximity matrix referred, this time, to a single configuration. Finally Umeyama (1988) came up with the definition of correspondence using the spectral decomposition of the adjacency matrices relative to weighed graphs of equal measure. The same author has later reconsidered (1993) these ideas, applying them for the solution of the matching problems among complex objects.

More recently many authors have tried to model the structural deformations of point sets. In this regard, Amit and Kong (1996) have used the graph theory to model the deformations of 2-D shapes contained in some medical images. Finally, other authors have formalised the concept of correspondence by using the links of a rigid body: we mention the work done by Morgera and Cheong (1995). Cross and Hancock (1988) proposed instead, for the correspondence problem solution, a statistical methodology based on the theorem of Bayes. Afterwards Luo

and Hancock (2002) have associated to this methodology a Procrustean criterion for the solution of an alignment problem between configurations with a different number of vertices.
Anyway, concerning the specific GIS and cartographic aspects, and the CAD/CAM related ones, the reported references do not provide a complete and satisfactory solution for the problem of identification of the correspondence between two or more sets of point coordinates. The above studies, in fact, have very general purposes, and give a correspondence solution also in the case of very dissimilar configurations: consequently, these methods are inadequate for the correspondence problem solution.

2. THE PROPOSED METHOD

The described procedure makes it possible to automatically recognise a given geometrical entity, represented by a finite number of vertices, into a more complex configuration, that can completely, or just partly, correspond with the entity taken into account. In more detail, the method allows both to identify and put in relation the single couples of homologous points of two representations of the same entity, and to establish the correspondence among points belonging to such entity and those ones belonging to the geometric configuration enclosing it entirely. For the two situations, the procedure just requires the knowledge of the vertex coordinates of the geometrical entities in the respective Cartesian reference systems. Therefore, the procedure can operate without the need that the points be acquired, or defined, according to a pre-fixed order, and also without the knowledge of any structural or topological information required to totally or partially characterise the connections among the various vertices.
The procedure foresees two distinct operational sequences according to the kind of the problem treated. In one case the process carries out the recognition of the homologous points of two correspondent representations of the same geometrical entity, in a very simple and direct way, independently of the assumed coordinate system and of the reciprocal scale factor. In the second case the process solves the inclusion problem by means of a more general methodology, capable to identify a geometrical configuration completely contained within another one, more complex and more extended, also in the case in which the approximate knowledge of the mutual scale factor is not available.
For both situations, the method is conceived on the use of mainly algebraic and geometric rules, and on the use of basic mathematical functions, chosen aiming to a fast and efficient software implementation.

2.1 The comparison case

The comparison problem considers generic point configurations, describing in the various cases, geometrical entities of the specific application fields (cartographic, CAD, GIS and so on). The comparison problem rises when:
- the geometrical entities considered are both composed of n points of known coordinates;
- the correspondence is bi-univocal: therefore each point of one configuration finds a correspondent in the other, and vice versa.

These conditions are necessary and sufficient: for the geometrical entities taken into account, further information is not required, like for instance the value of the mutual scale factor or the criterion by which the points are listed. For this procedure, each entity is represented by a geometrical configuration of points, without any information of structural or topological kind; every configuration is uniquely described by its vertex coordinates in a proper and independent 2D or 3D Cartesian reference system. If the geometrical configurations do not have axes or planes of symmetry, for which not unique solutions might happen, the procedure makes it possible to specify the correspondences between homologous points, also in the presence of errors in the coordinate values of the considered vertices.
Before describing in detail every phase, it is necessary to premise some definitions, useful in the following explanations.
Let us define "*rigid link*" the segment joining two vertices (u, v), arbitrarily chosen, of a same configuration C. If C is characterised by n vertices, it follows that every vertex v belonging to C has n-1 rigid links defined with respect to the remaining vertices. Let us write as l_{iv} the i^{th} rigid link referred to the vertex v, and also the length of the same link.
Let us call "*ordered set of rigid links*" of the vertex v the following set: $I_v = \{l_{1v}, l_{2v}, \ldots, l_{(n-1)v}\}$ with $l_{iv} \le l_{(i+1)v}$, and with $i = 1, \ldots, n-2$.
Let u and v be two vertices belonging to a generic configuration C of n points; let I_u e I_v be the respective ordered sets of the rigid links. Furthermore, let us define "*homologous rigid links unconformity*" of the vertices u and v the following value:

$$\lambda_{uv} = \left|I_v - I_u\right| = \sum_{i=1}^{n-1}\left|l_{iv} - l_{iu}\right| .$$

with $l_{iv} \in I_v$ and $l_{iu} \in I_u$

To each vertex v of a configuration C it can be associated a "*set of asymmetry distances*" that assumes the following form:

$$D_v = \left\{d_{1v}, d_{2v}, \ldots, d_{(n-2)v}\right\}$$

with $d_{iv} = l_{(i+1)v} - l_{iv};\ l_{iv} \in I_v;\ i = 1, \ldots, n-2.$

2.1.1 The direct method: Once defined the necessary tools, let us pass to the procedure description. The only available data for solving the comparison problem are the vertex coordinates characterising the geometrical configurations A and B, listed in the respective and homonymous matrices.
An initial test tries to single out possible symmetrical axes. If symmetrical configurations are present, the comparison problem will be solved by the alternative general procedure, since there exists the possibility of not univocal solutions.
The successive phase foresees the computation of the unknown scale factor. The estimate of the scale factor s that joins **B** to **A** is given by:

$$s = \sqrt{\frac{\mathrm{trace}\left(\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A}\right)}{\mathrm{trace}\left(\mathbf{B}^{\mathrm{T}}\mathbf{W}\mathbf{B}\right)}} . \tag{1}$$

where $\mathbf{W} = \mathbf{I}\text{-}\mathbf{j}^{\mathrm{T}}\mathbf{j}/\mathbf{j}\mathbf{j}^{\mathrm{T}}$ is applied to centre the original point coordinate values into their respective systems, $\mathbf{j} = [1, 1, \ldots 1]$ is the unit vector of dimension equal to the number of points contained in **A** (or **B**), and **I** is the identity matrix.
The scale factor can be removed pre-multiplying the point coordinates of the configuration B by the scalar s.
The next step is to look for a particular vertex in one of the two configurations, for example in A. This is named "*point of maximal asymmetry*", and identified with the symbol m_A. Since the coordinate values are affected by errors, it is necessary to fix a "significant" threshold (Δ); the value of Δ, computed according to the length of the rigid links present in A, depends on the fixed tolerance values, or on the entity of the errors characterising the coordinates.
Now let v be any vertex of A. Let us call "*index of significance of v*" ($i_{a\text{-}v}$) the number of elements belonging to D_v having a

value greater than the significant threshold. This corresponds to find out in I_v the number of lengths l_{iv} significantly different among all of them, without the uncertainty due to the coordinate errors. The vertex with the maximal asymmetry will correspond to that particular point of A having the largest index of significance. Formally, m_A is such that:

$$\mathrm{i}_{a-m_{\mathrm{A}}} = \max_{v \in \mathbf{A}} \left(\mathrm{i}_{a-v} \right). \tag{2}$$

In case of ambiguity, due to the simultaneous presence of more vertices having the same maximal value of reliability, the vertex that, for the same i_a, is characterised by the minimal component of D_v with the largest value, will be chosen as the point of maximal asymmetry. The reason and the convenience of finding out a point of maximal asymmetry, will be clear in the following.

At the next step, the corresponding point of m_A in the configuration B is researched; this vertex is indicated with the symbol m_B. To this purpose, the necessary but not sufficient condition to state that two arbitrary vertices (u and v) are correspondent, for configurations with the same scale rate, is that, for each rigid link associated to u, there must exist another one of the same length, a part for random errors associated to v.

Referring to this property and considering the fact that the symmetrical configurations have been excluded, the most probable correspondent of m_A is, among all the vertices of B, that one with the less discrepancy to homologous rigid links computed with respect to the point of maximal asymmetry in A. Formally, m_B is such that:

$$\lambda_{m_{\mathrm{A}}-m_{\mathrm{B}}} = \left| \mathrm{I}_{m_{\mathrm{A}}} - \mathrm{I}_{m_{\mathrm{B}}} \right| = \min_{v \in \mathbf{B}} \left(\left| \mathrm{I}_{m_{\mathrm{A}}} - \mathrm{I}_v \right| \right) \tag{3}$$

In the case of exactly comparable configurations, that is without errors in the coordinate values, the discrepancy to homologous rigid links between two corresponding points is necessarily equal to zero. All this leads to an important consequence: the comparison between λ and a proper threshold (L), proportional to the admitted tolerances, provides a criterion to evaluate the correspondences. In fact, given two points u and v, if $\lambda > L$, the two considered vertices are not correspondent. Given the characteristic function assumed, L is defined as "*correspondence threshold*".

Keeping in mind what already exposed, let us consider again the points of maximal asymmetry and their correspondence ratio. The verification that m_A e m_B are not correspondent vertices, leads to consequences much more significant than a wrong correspondence. Referring to the way in which m_B has been determined, a negative exit of the mentioned test would imply the conclusion of the procedure: which other point in B could overcome the comparison?

This impossibility to proceed is due to the fact that the comparability hypothesis between the configuration considered it is not verified. A positive exit of the test about the correspondence between m_A and m_B allows, on the contrary, to pass to the next phase.

Now, a fast and tentative problem solution of the residual correspondences will be looked for. The final solution should try to satisfy the following aspects: to be, as much as possible, near to the real result, and to require a light computational work.

Now let u be a generic vertex of A, and v be the correspondent one in B: the method can solve the residual correspondences between A and B by the comparison of the rigid link lengths of u and v assumed as reference points.

Let a_i be the vertex of A joined to u by the rigid link l_{iu} of I_u. Let b_i be the vertex of B joined to v by the rigid link l_{iv} di I_v. Since the correspondent vertices define, between A and B, rigid links of the same length, a part for some random errors, and remembering the definition of "*ordered set of rigid links*", it seems correct the hypothesis that the correspondent vertex of a_i is b_i, with $i = 1, ..., n-1$. This assumption is as much close to the real situation as much the correspondence between the rigid links of the reference points is univocally identifiable. It is evident that the presence of links with very similar length (if not equal), combined to the distortions caused by the errors, can lead to imprecise results in the proposed solution.

With the aim to optimise the procedure, it is therefore necessary to identify the reference vertex of A providing the maximal reliability and correctness to the hypothesised solution. This is the vertex whose rigid links have as much as possible different lengths with respect to the other points. According to its definition, the reference point for the correspondence solution is the point of maximal asymmetry. If m_A is known, it is possible to find its correspondent m_B, and the residual correspondences can be immediately identified at the first tentative.

In the following, the quality of the obtained solution is evaluated. First of all it is necessary to define which, among the proposed correspondences, can originate doubts about their effective correctness. As mentioned before, these will be the correspondences characterised by rigid links, connected to m_A, having similar length, that is contained within the tolerance range.

According to what already explained, the methodology to search for couples of vertices with a doubt correspondence, assume the following form: fixed the i^{th} distance of asymmetry: $d_{im_{\mathrm{A}}}$ with $d_{im_{\mathrm{A}}} \in \mathrm{D}_{m_{\mathrm{A}}}$; if $\Delta > d_{im_{\mathrm{A}}}$; then $(a_{i+1}; b_{i+1})$ and $(a_i; b_i)$ are doubt correspondences.

The same reasoning is repeated for all the distances of asymmetry referred to the vertex m_A. At the end of this step, if the test has not found doubt correspondences, the problem can formally be considered solved.

2.1.2 Validation test: To make this method satisfying the maximal reliability, the true final step is to verify the hypothesis of comparability of the configurations A and B. Having identified all the correspondences is not, by itself, index of correctness; the method of research, based on the lengths of the rigid links, leaves out of consideration from the effective spatial disposal of the vertices. To confirm the supposed comparability it is necessary to verify that the considered configurations have, a part for some random errors, the same shape. We can state that two configurations have the same shape if they can be put coincident by rotations, translations and isotropic deformations.

Let us indicate with e^2 the square mean of the measured distances among correspondent points belonging, from one side, to A (reference configuration) and, from the other, to a generic configuration having the same shape of B (transforming configuration). To perform the so called "test of shape", the first step is to compute the minimum value that can be assumed by e^2; this value is reported with the symbol ε^2.

Given:

$\mathbf{A} = \{x_{A1}, x_{A2}, .., x_{An}\}$ = Reference configuration

$\mathbf{B} = \{x_{B1}, x_{B2}, .., x_{Bn}\}$ = Transforming configuration;

the term ε^2 is provided by the following equation (Umeyama, 1991):

$$\varepsilon^2 = \sigma_{\mathrm{A}}^2 - \frac{\mathrm{tr}\left(\mathbf{DS}\right)}{\sigma_{\mathrm{B}}^2} \tag{4}$$

where:

$$
\begin{aligned}
&\mathbf{m}_{\mathrm{A}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{\mathrm{A}i} \quad \sigma_{\mathrm{A}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left\|\mathbf{x}_{\mathrm{A}i} - \mathbf{m}_{\mathrm{A}}\right\|^2 \\
&\mathbf{m}_{\mathrm{B}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{\mathrm{B}i} \quad \sigma_{\mathrm{B}}^2 = \frac{1}{n}\sum_{i=1}^{n}\left\|\mathbf{x}_{\mathrm{B}i} - \mathbf{m}_{\mathrm{B}}\right\|^2 \\
&\mathbf{K}_{\mathrm{BA}} = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbf{x}_{\mathrm{B}i} - \mathbf{m}_{\mathrm{B}}\right)\left(\mathbf{x}_{\mathrm{A}i} - \mathbf{m}_{\mathrm{A}}\right)^{\mathrm{T}} = \mathbf{UDV}^{\mathrm{T}} \quad \text{(SVD)} \\
&\mathbf{S} = \begin{cases} \mathbf{I} & \text{if } \det\left(\mathbf{K}_{\mathrm{BA}}\right) \geq 0 \\ (1,1,\ldots 1,-1) & \text{if } \det\left(\mathbf{K}_{\mathrm{BA}}\right) < 0 \end{cases}
\end{aligned}
\tag{5}
$$

In this formulation, the vertices of the transforming configuration correspond to those ones of the reference configuration. It happens therefore that x_{Ai} corresponds to x_{Bi}, with i = 1, …, n.. Furthermore, in the Singular Value Decomposition (SVD) of $\mathbf{K}_{BA}$, the eigenvalues have positive values, and are located in $\mathbf{D}$ in decreasing order ($\mathbf{D} = \mathrm{diag}(d_j)$; $d_1 \geq d_2 \geq \ldots \geq d_k \geq 0$ where k is the dimension of the reference system).

The term ε^2 provides an index of the shape difference for the two considered configurations ($\varepsilon^2 = 0$ means that the two configurations have the same shape). To limit the shape difference to the only presence of measurement random errors, it is necessary that the range of the values the term ε^2 can assume must be contained within a given threshold. In the proposed procedure this threshold will be represented by a shape parameter δ^2 referred to the reference configuration. The test of shape is accepted for $\varepsilon^2 < \delta^2$: this event definitely ends the comparison procedure.

If this does not happen, and previously the presence of doubt correspondences had been recognised, it is necessary to identify which ones of the correspondences are wrong or are inverted.

Remembering what already said about the correspondent points and the rigid links, the determination is based on the following criterion: for each doubt correspondence (a_i; b_i); if $\lambda_{a_i - b_i} > \mathrm{L}$; then ($a_i$; b_i) is a wrong correspondence. The same reasoning is repeated for all the doubt correspondences found out at the preceding step. At the end of this phase, three are the cases that can be present:

- there is only one wrong correspondence;
- the number of the wrong correspondences is equal to two;
- the number of the wrong correspondences is greater than two.

The first situation means, for the largest part of the cases, that all the proposed correspondences are apparently correct: the final decision is then remitted to the shape test. If the number of the wrong correspondences is equal to two, a swap is carried out: if (a_i; b_i) and (a_j; b_j) correspond to the wrong correspondences, then (a_i; b_j) and (a_j; b_i) represent the necessary correction. In the third case, the most complex, the procedure is iterated just for the subsets of A and B containing correspondences not yet solved, recalling again the considerations done for the two preceding cases.

An alternative solution of the comparison problem can be found in Sossai (2003), and in Beinat, Crosilla & Sossai (2003, 2004).

2.2 The inclusion case

The situation of inclusion occurs when a group of points is entirely contained in a more numerous and topographically extended set. This means that the geometrical entities taken into consideration can have a different number of points: the configuration with less vertices will be called "enclosed", while the other will be the "enclosing" one. This situation is not mandatory: the general method proposed works well also for configurations having the same number of points, solving in this way, by another approach, the comparison case.

Other conditions are required instead: all the points of the enclosed configuration must find univocal correspondence within the enclosing configuration, and every entity must not be represented by a degenerate geometrical configuration, like that whose vertices are all approximately aligned. As for the preceding problem, also for this case, we do not need the knowledge of the scale ratio between the configurations, neither any structural or topological information.

According to these conditions, the method described in the following allows to identify the correspondences between homologous points, also in the presence of some errors present in the coordinates of the considered vertices. In Figure 1 and Figure 2 two different problems, relative to two distinct application fields, are shown. In the following, we will indicate with A the enclosed configuration, and with B the enclosing one.

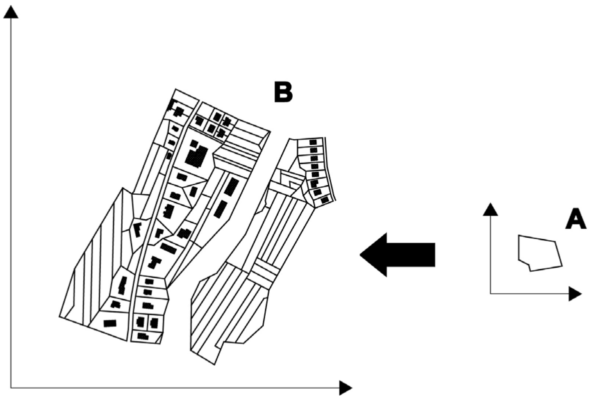

Figure 1. The problem of inclusion for a cadastral map, that is to find the point correspondences between a parcel A fully contained into a more general map B.

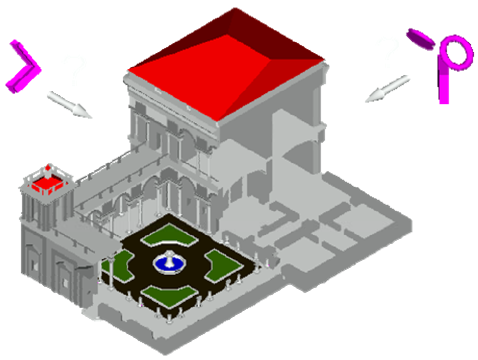

Figure 1. The problem of inclusion for a CAD design, that is to locate some predefined structural elements fully contained into a more complex drawing.

2.2.1 The general method using a correspondence kernel:

The idea leading to the general solution of the problem of correspondence is based on the following two considerations:

- the less is the number of the vertices of the enclosed set (considering constant the number of the enclosing points) the easier is the solution of the inclusion problem;

- the correspondence problem solution valid for an entire enclosed configuration, must also be appropriate for every partial configuration of it.
It is therefore valid and correct to devote the attention, not to the entire enclosed configuration, but to a particular and characteristic minimal subset of it, constituted by three vertices, called "*basic triangle*" or also "*correspondence kernel*".
Once such a triangle in the enclosed configuration is identified, its images within the enclosing configuration have to be looked for. To do so, the basic triangle is overlapped to each image by a Procrustean similarity transformation, and the condition that every point of the remaining configuration A has a correspondent one in the configuration B is verified. The final solution is that one furnishing the best geometrical fit, and the correct correspondence for all the points of A.
This procedure is described in the following (see for more details: Sossai, 2003; Beinat, Crosilla & Sossai, 2004).

2.2.2 Construction of the basic triangle: The process starts finding out, within the enclosed configuration A, the two vertices separated by the largest distance. Let us call 1 and 2 these points, name *a* the segment 1-2, and d_a its length. One of the points, for instance point 2, it is assumed as the reference .
The third vertex is chosen, among the remaining points, as the closest point to the reference. Called *v* the possible candidate, the distances d_{v1} e d_{v2}, between *v* and the already defined points 1 and 2, are calculated. In order that *v* can be considered the third vertex of the basic triangle, the following conditions must be verified:
- $||d_{v1} - d_{v2}| - d_{12}| \geq \Delta_0$;
- $|d_{v1} - d_{v2}| \geq \Delta_1$, $|d_{v2} - d_a| \geq \Delta_1$, $|d_a - d_{v1}| \geq \Delta_1$.

The symbols Δ_0 e Δ_1 are threshold values proportional to the required tolerances. The first condition aims to avoid that the basic triangle be degenerate, for instance too flat. The second condition guarantees that the basic triangle is not characterised by symmetrical axes, that is, neither isosceles, nor equilateral. These constraints are chosen in order to define one configuration able to minimise the ambiguity and the possibility of errors, both in the phase of identification that during the solution of the correspondences.
Let us call point 3 such a vertex and name with *b* the segment 1-3, with *c* the segment 2-3 (the shortest) and with d_b e d_c their respective lengths. The configuration 1-2-3, so determined, represents the basic triangle.
Since we are using coordinate values affected by errors, a variability interval of the distances between the vertices must be considered. For this reason the lengths of the sides d_a, d_b and d_c of the original triangle are substituted by the variability intervals so defined:

$$R_b = [d_b\text{-t}; d_b\text{+t}] \qquad (6)$$
$$R_c = [d_c\text{-t}; d_c\text{+t}]$$

where t is the tolerance parameter, that can be fixed or proportional to the distance d_{ij}. If the scale factor *s* in known, the procedure defines also the rounded value $R_a = [d_a\text{-t}; d_a\text{+t}]$. As a last thing, a shape parameter δ^2 is computed, whose value is proportional to the square mean of the tolerances computed for the centroid distances of the correspondence kernel.

2.2.3 Images of the basic triangle: Once the basic triangle and its admitted interval of variability are fixed, its possible images in B are looked for. This generates a set of point triplets of possible correspondence.
For the most general conditions, one triplet of possible correspondence is a general subset of three vertices [*i*, *j*, *k*] of B not degenerate, characterised by the fact that $d_{ij} < d_{jk} < d_{ki}$, where the symbols identify the distances between the points *i* and *j*, *j* and *k, k* and *i*, respectively.
For the definition of the point triplets of possible correspondences two alternatives can be present. If the scale factor *s* is unknown, all the triplets of possible correspondence will belong to the above mentioned set, and it will be necessary to proceed with a combinatorial approach. In the opposite case, if the scale factor is known, just the triplets of possible correspondence, such that $s \cdot d_{ki}$ is contained within the round R_a, will belong to the set. Once this particular set is constructed, the attention moves to individuate, within the same set, the subset of the kernels of possible correspondence.
The kernels for which exists a relationship of comparison with the basic triangle, will belong to this set, while all the others are excluded. Let us consider, for example, a generic triplet of points [*i*, *j*, *k*] (note that the meaning of the symbols previously introduced remain valid). It follows a geometric comparison, based on the distances, between the triplet of points of possible correspondence and the basic triangle. If the scale factor s is not known, this will be estimated as a/d_{ki}.
For $d_b > d_c$ (case 1) it is:
segment *i-j* = possible image of the segment *c*;
segment *j-k* = possible image of the segment *b*;
segment *k-i* = possible image of the segment *a*.
In this case, if $s \cdot d_{ij}$ belongs to the round value R_c and if $s \cdot d_{jk}$ belongs to the round value R_b, a relation of comparison exists.
For $d_b < d_c$ (case 2), *i-j* and *j-k* are swapped, so to correspond to possible images of *b* and *c* respectively. If the comparison fails, the examined triplet of points will be rejected, and the process passes to examine the successive.
In the comparison test, the correspondence among the points *i*, *j*, *k* of the triplet, and the vertices 1, 2, 3 of the basic triangle, is immediately defined. They remain valid, in fact, the following results:
in the case 1, $i \equiv 2$; $j \equiv 3$; $k \equiv 1$;
in the case 2: $k \equiv 1$; $j \equiv 3$; $i \equiv 2$.
The identification of a kernel of possible correspondence is completed with the execution of the test of shape, already introduced in the previous chapter. If also this test is overcome, it follows the expansion of the kernel by the solution, where possible, of the residual correspondences till the identification of a possible complete image of the enclosed configuration.

2.2.4 The expansion of the correspondence kernel: During this phase, the basic triangle is overlapped to all the selected correspondence kernels, by an appropriate algorithm, obtained from the orthogonal Procrustes analysis. This technique allows to obtain the mutual least squares fit of a moving matrix configuration with respect to another one, considered as a reference, by means of a proper set of transformation parameters, so to satisfy a prefixed objective function.
In detail, the Procrustes algorithm makes it possible to find out the rotation matrix **R**, the translation vector **t**, and the isotropic deformation *s*, to apply to the moving configuration, to satisfy the minimum to the distance square mean ε^2 among the correspondent points belonging to the two considered configurations (see e.g. Beinat & Crosilla, 2003a).
Let:

Y = $\{y_1, y_2, .., y_n\}$ be the reference configuration, that is the possible correspondence kernel;
X = $\{x_1, x_2, .., x_n\}$ be the moving configuration, that is the basic triangle.

Referring to the notation previously introduced, the solution proposed by Umeyama (1991) is now explained. In the case in which the rank of K_{xy} is equal to k, it follows:

$$\mathbf{R} = \mathbf{USV}^{\mathrm{T}} \tag{7}$$

$$\mathbf{t} = \mathbf{m}_{\mathrm{y}} - s\mathbf{Rm}_{\mathrm{x}} \tag{8}$$

$$s = \frac{1}{n}\mathrm{tr}\left(\mathbf{DS}\right) \tag{9}$$

if the rank of K_{xy} is equal to k-1, the elements of matrix **S** in the SVD of **R** are the following:

$$\mathbf{S} = \begin{cases} \mathbf{I} & \text{if } \det(\mathbf{U})\det(\mathbf{V}) = 1 \\ \mathrm{diag}(1,1,\ldots 1,\text{-}1) & \text{if } \det(\mathbf{U})\det(\mathbf{V}) = -1 \end{cases} \tag{10}$$

The basic triangle will be therefore rotated, translated and scaled, according to the rules of the Procrustes analysis, to fit in the best way the kernel of possible correspondence. The application of the just mentioned transformation parameters to the entire configuration, allows, with good approximation, the insertion of this configuration in the datum of the enclosing configuration.

Once the two geometrical entities are represented in the same reference system, the solution of the residual correspondences is obtained by a simple comparison of the mutual distances. The nearest point B to a fixed vertex of A (expressed in the new coordinate system) will be its probable correspondent.

By solving all the residual correspondences, one possible complete image of the enclosed configuration A in the enclosing B, can be identified.

The final solution is defined, also in this case, by a test of shape. For each possible image previously identified, the value of the Procrustean shape parameter ε^2 is determined, assuming as reference the enclosed configuration. The most probable image will be that one corresponding to the lowest value of ε^2.

3. CONCLUSIONS

In this paper we have illustrated a novel procedure to automatically identify correspondences between geometrical entities missing of topological structure. The method is based only on the knowledge of the vertices coordinates describing the geometrical entities considered, referred to their own and different Cartesian reference systems.

The entire process is developed without the necessity that the points are acquired or defined according to a prefixed order, and without requirements about structural or topological information, relative to the links among such a vertices. The recognition of the homologous points of two correspondent configurations of the same geometrical entity is solved, independently for the coordinate systems assumed, and for the reciprocal scale rate.

For the inclusion problem solution, the methodology is able to identify a geometrical configuration completely contained within another one, more general, also in the case in which the approximate knowledge of the reciprocal scale factor is not available.

Further developments of the proposed method will consider the case of partial inclusion, that is the identification of the subset of common points belonging to two different configurations. Finally, a research field will be the identification of shape parameters alternative to that employed, and the introduction of methods and principles of the fuzzy logic.

AKNOWLEDGMENTS

This work was carried out within the research activities supported by the INTERREG IIIA Italy-Slovenia 2003-2006 project "Cadastral map updating and regional technical map integration for the Geographical Information Systems of the regional agencies by testing advanced and innovative survey techniques"

REFERENCES

Amit, Y., Kong, A., 1996. Graphical templates for model registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18, pp. 225–236

Beinat, A., Crosilla, F., 2003. Generalised Procrustes Algorithms for the Conformal Updating of a Cadastral Map. *Zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 5, pp 341–349.

Beinat, A., Crosilla, F., Sossai, E. 2003. Ricerca automatica di corrispondenze fra entità geometriche di una cartografia catastale. *Atti della 7a Conferenza Nazionale ASITA*, Verona, 28-21/11, pp. 241–246.

Beinat, A., Crosilla, F., Sossai, E., 2004. Riconoscimento automatico di entità geometriche non strutturate di una cartografia catastale. *Bollettino della SIFET*, *(in printing).*

Cross, A.D.J., Hancock, E.R., 1988. Graph matching with a dual-step EM algorithm. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(11), pp. 1236–1253.

Luo, B., Hancock, E., 2002. Iterative Procrustes alignment with the EM algorithm. *Image and Vision Computing*, 20, pp. 377–396.

Morgera, S.D., Cheong, P.L.C., 1995. Rigid-body constrained noisy point pattern-matching. *IEEE Transactions on Image Processing*, 4(5), pp. 630-641.

Papadimitriou, C. H., Steiglitz, K., 1982. *Combinatorial optimisation algorithm and complexity*, Prentice-Hall, Englewood Cliffs, NJ.

Scott, G.L., Longuet-Higgins, H.C., 1991. An algorithm for associating the features of 2 images. *Proc. of the Royal Society of London Series B (Biological)*, 244(1309), pp. 21–26

Shapiro, L.S., Brady, J.M., 1992. Feature-based correspondence – an eigenvector approach. *Image and Vision Computing*, 10, pp. 283–288.

Sossai, E., 2003. *Ricerca automatica di corrispondenze fra entità geometriche di una cartografia catastale*. Degree Thesis. Faculty of Engineering. University of Udine. Academic Year 2002-2003.

Umeyama, S., 1988. An eigendecomposition approach to weighted graph matching problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10, pp. 695–703.

Umeyama, S., 1991. Least squares estimation of transformation parameters between two point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4), pp. 376-380.

Umeyama, S., 1993. Parameterised point pattern matching and its application to recognition of object families. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15, pp. 136–144.

Ullman, S., 1979. *The Interpretation of Visual Motion*. MIT Press, Cambridge, MA.