ADAPTIVE TRANSFORMATION OF CARTOGRAPHIC BASES BY MEANS OF MULTIRESOLUTION SPLINE INTERPOLATION

Maria Antonia Brovelli, Giorgio Zamboni

Politecnico di Milano – Polo Regionale di Como, Via Valleggio 11 – 22100 Como
+390313327517, fax +390313327519, e-mail maria.brovelli@polimi.it
+390313327528, fax +390313327519, e-mail giorgio.zamboni@polimi.it

KEYWORDS: cartography, GIS, integration, algorithms, multiresolution, vector

ABSTRACT:

GIS databases often need to include maps from diverse sources. These can differ one another by many characteristics: different projections or reference systems, (slightly) different scales, etc. Theoretical and/or empirical transformations are available in literature to obtain maps in a unique system with a fixed tolerance. These transformations are nevertheless insufficient to completely remove differences and deformations: the outcome is that the geographic features on the maps do not fit in a perfect way. To reduce the deformation several transformations (affine, polynomial, rubber-sheeting) exist. The paper presents a new approach to the problem based on an interpolation by means of multiresolution spline functions and least squares adjustment. One map is taken as reference and the others are warped to comply with it. The interpolation is made by comparison of coordinates of a set of homologous points identified on the maps. The use of spline functions, compared to affine or polynomial interpolation, allows to have a greater number of coefficients to make more adaptive and localized the transformation. The multiresolution approach removes the rank deficiency problem that ordinary spline approach suffers for. Moreover the resolution of the spline functions depends areawise on the spatial density of homologous points: the denser are the points in the area, the better adapted to them can be the interpolating surface. A statistical test has been built to automatically choose the maximum exploitable resolution. The paper presents the method and one application in the example.

1. INTRODUCTION

1.1 Interoperability in Geographic Information Systems

The increase of application fields of GIS (local administration, tourism, archaeology, geology, etc.) has made of new interest the study of the sharing of information from different geographic databases, also known as "GIS data interoperability".
In general, with the technical term interoperability we define a user's or a device's ability to access a variety of heterogeneous resources by means of a single, unchanging operational interface. In the GIS domain, interoperability is defined as the ability to access multiple, heterogeneous maps and corresponding geo-referenced data (either local or remote) by means of a single, unchanging software interface.
Interoperability engages at several levels: network protocol, hardware & OS, data files, DBMS, data model and application semantics. Nowadays greater automation is already evident, especially at the first four levels of interoperability; however at the most fundamental levels (data model and semantics) there remains further room for improvement.
Usually geographic information is formed by geometric and thematic attributes. For this reason the research on interoperability is focused on topological compatibility (at the level of data structure) and on semantic compatibility (at the level of identifiers) of the data.
To guarantee the interoperability there is another very important problem often not mentioned: the geometrical compatibility (at the level of coordinates) of the maps.
GIS databases often include maps coming from diverse sources. These can differ one another by many characteristics: different projections or reference systems, (slightly) different scales, different kinds of representations, etc., with the result of geometric incompatibility of the different maps.

1.2 The "Conflation Maps" problem

Map conflation was first addressed in the mid-1980s in a project to consolidate the digital vector maps of two different organizations (Saalfeld, 1988). The problem was split into two parts: the detecting of homologous elements between the two maps, and the transformation of one map compared with the other (Gillman 1985; Gabay and Doytsher, 1994). Point elements within one map were selected as the group of features whose counterpart points on the other map enable the conflation process (Rosen and Saalfeld, 1985).
Since then, many conflation algorithms have been developed and improved. Recently, the main concern has been focused on data integration. Several geodata sets which cover the same area but are from different data providers, may have different representation of information and may be of different accuracy and forms.
Conflation can be used to solve different practical problems like spatial discrepancy elimination (such as sliver polygons, shifts of features, etc.), spatial feature transfer (new features can be added into the old map, or old coordinates can be update), attribute transfer (i.e. the attributes in the old maps can be transferred into the new maps).
The conflation algorithms can be classified into three kinds: geometric, topological and attribute method.
Geometric methods are mostly used because we are dealing with spatial objects. They scan geometric objects from both data sets and compare them by geometrical criteria: distance,

angular information of linear objects, location relationships, shape feature of the objects (i.e. lines length, polygon perimeter and area, etc.). The geometric method is used in most cases and requires that two data sets have similarity in geometric location, thus map alignment or rubber sheeting may be involved in the processing.
Topological methods use topological information such as connectivity between lines, adjacency between polygons and composition relationships to correlate objects: arcs meet at a node, arcs form a polygon, and so on.
Topological matching is usually used to reduce the search range or check the results of geometric matching and it can be used only when topological information is available.
The attribute method is also referred to as the semantic method. This method can be used to match features if both data have common attributes with the same known semantics. Otherwise a relationship table must be established.
Once the correspondence between different data sets are established, the spatial features need to be put together and some transformation may be done so that the data describing the same object coincide.
Theoretical and/or empirical transformations are available in literature to obtain maps in a unique system with a fixed tolerance. These transformations are nevertheless insufficient to completely remove differences and deformations.

1.3 Problems of the most common transformations

Geometric transformation is the process of converting a digital map from one coordinate system to another by using a set of control points (also known as homologous pairs) and some transformation equations. There are several types of transformations.
Polynomial transformations between two coordinate systems are typically applied in cases where one or both of the coordinate systems exhibit lack of homogeneity in orientation and scale. The small distortions are then approximated by polynomial functions in latitude and longitude or in easting and northing. Depending on the degree of variability in the distortions, approximations are carried out using second, third, or higher degree polynomials.
Polynomial approximation functions themselves are subject to variations, as different approximation characteristics may be achieved by different polynomial functions. The simplest polynomial is the affine transformation (or first order polynomial) which is a 6 parameter transformation (rotation, shift in X and Y, differential scaling in X and Y and skew).
The most important advantages using affine transformation are: straight lines are transformed in straight lines; parallel lines are transformed in parallel lines; incident lines are transformed in incident lines; the ratio between parallel lines is preserved.
Using particular restriction on the polynomial coefficients it is also possible to preserve the areas of the features (congruence transformation) or the shapes (similarity transformation).
The higher order polynomials are useful in registering maps with varying localized distortions, i.e. where the distortion can not be easily modelled by affine transformation. The greater the local distortions is, the higher is the polynomial function to be used.
One property of polynomial interpolation, which is undesirable in practical applications, is that the interpolating polynomial can show oscillations which are not present in the data. These oscillations get worse as the degree of the polynomial increases. To clarify this concept a one dimensional case can be taken into account: figure 1 shows the famous example of this phenomenon due to Runge using 11 equally spaced data points on the interval [-1,1] and the interpolating polynomials of different degree (3,7 and 11 respectively).

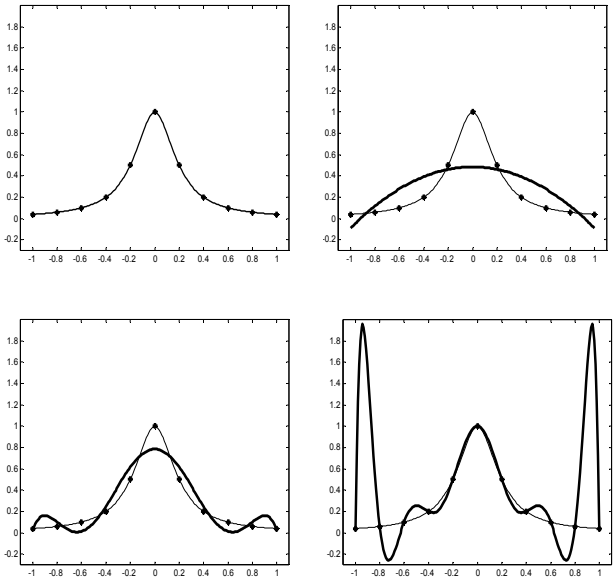


Figure 1. Oscillation problem of polynomial interpolation

The second method commonly used applies a variable transformation to different portions of the unadjusted data. A possible solution is based on the triangulated data structures method suggested by Gillman (1985) and Saalfeld (1987) and a piecewise linear homeomorphic transformation, known also as rubber sheeting, suggested by White and Griffin (1985), Saalfeld (1985) and Gabay and Doytsher (1995). This approach, again based on homologous points of the two maps, is today the most popular (Lupien and Moreland, 1987; Doytsher and Hall, 1997; Cobb et al., 1998).
The main disadvantage of the rubber sheeting transformation is that it holds the control points fixed, that is the control points in the two maps match precisely, therefore they are treated as being completely known and with no error. This kind of approach is purely deterministic and it doesn't consider the fact that any coordinate in a geographic database has a measurement error. This second consideration is particularly important: while rubber sheeting allows for a better solution from the numerical point of view (the control points coincide and therefore null residuals are obtained) it can bring about the description of the phenomenon of transformation far from the physical reality.
Another problem related to the rubber sheeting transformation is that each error in the selection of control points affects, without any error filtering, the deformation of its no homologous neighbouring points.

2. THE INTEGRATION PROCESS

2.1 The first step: the automatic research of homologous points

The starting point to estimate every transformation is the homologous points detection.
Control points are points that are in the same location in both datasets. Usually they are manually chosen interactively in

both datasets: they are displayed on the screen and the user clicks on a location in one map and then the same corresponding location in the second map. Typically control points are easily identifiable features such as building corners, major road intersections, etc. The selection of control points must be done carefully; their number and the quality influences the types of curve fitting that can be performed (i.e. at least four points are needed for an affine transformation estimation). Moreover control points must be spatially scattered over the datasets and in a number greater than the minimum necessary to compute the parameters of the chosen transformation.
The estimate of the parameters, independently from the kind of transformation used, becomes better (more accurate) as the number of homologous points increases.
To avoid the time-consuming manual search of these correspondence and the possible human errors, a strategy is needed to automate the procedure.
The idea is to reproduce as much as possible what the operators manually do when they try to superimpose two maps: they visually search for the same geographic features represented on the two different cartographic supports. To detect the feature related to a certain entity the operators implicitly makes at the same time geometric, semantic and topological analyses.
During the visual analysis, the operators compare the shape of the features on the maps. We can summarize this operation by considering three steps: an analysis of the coordinates of the points that geographically describe the shape of the objects, an analysis of the "directional" compatibility of the segments starting from the points and finally a semantic analysis.
Therefore, the basic hypothesis is that, since every cartographic entity is essentially defined by points (coordinates) and semantic attributes, the simplest way to make the search is to focus on them: a point P_1 on map c_1 is homologous of a point P_2 on map c_2 if the geographic feature related to the two points corresponds: figure 2 shows the example of homologous points that can be manually detected on two corresponding maps.

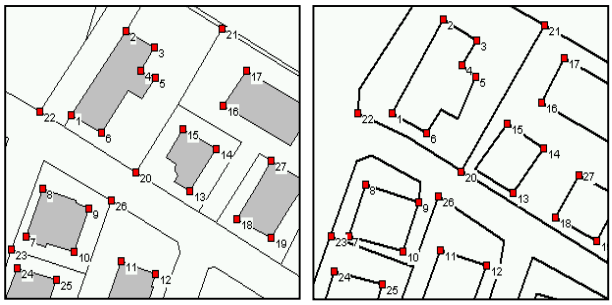


Figure 2. Homologous points on two different maps

2.2 The second step: the choose of the transformation

Once homologous pairs have been detected a warping transformation follows to optimally conflate the different maps. To make it more adaptive and localized a combination of finite support functions can be used.
In this way the estimation of each function coefficient will only depend on the data within the corresponding finite domain. The most common functions used for this estimation approach are the splines.

2.2.1 The classic spline interpolation approach

In general terms, we want to interpolate a field $d(\underline{t})$ sampled on N spread points $\underline{t}_1, \underline{t}_2, \ldots, \underline{t}_N$ in a plane.
The main idea is that the observed value $d_o(\underline{t})$ can be modelled by means of opportune spline combinations (deterministic model) and residuals $\underline{v}_i$ thought as noises (stochastic model).
The one-dimensional 0 order spline (see figure 3.a) is defined as:

$$\varphi^{(0)}(t) = \chi_{[0,1]}(t) = \begin{cases} 1 & t \in [0,1] \\ 0 & t \notin [0,1] \end{cases} \quad (1)$$

The basic function can be shifted and scaled throw:

$$\varphi_{jk}^{(0)}(t) = \varphi^{(0)}(2^j t - k) \quad (2)$$

where j fixes the scale and k the translation.
The splines of higher orders can be obtained starting from $\varphi^{(0)}(t)$ by means of convolution products.
The expression of the first order mono-dimensional spline (see figure 3.b) is therefore:

$$\varphi^{(1)}(t) = \varphi^{(0)} * \varphi^{(0)}(t+1) \quad (3)$$

and then:

$$\varphi^{(1)}(t) = \begin{cases} 1-t & 0 \le t \le 1 \\ 0 & t > 1 \\ \varphi^{(1)}(-t) & t < 0 \end{cases} \quad (4)$$

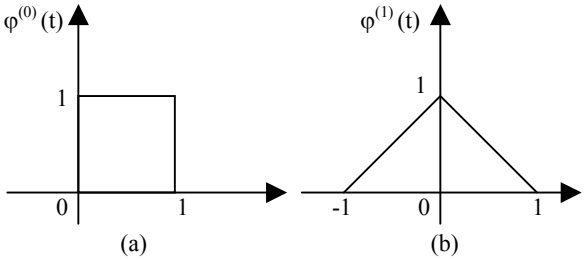


Figure 3. Mono-dimensional 0 (a) and 1 (b) order splines

Using a linear combination of g order splines with a fixed j resolution we obtain the following function:

$$d(t) = \sum_{-\infty}^{+\infty} {}_k \lambda_k^{(g)} \cdot \varphi^{(g)}(2^j t - k) \quad (5)$$

which represents a piecewise polynomial function on a regular grid with basic step $[k2^{-j}, (k+1)2^{-j}]$.

The corresponding bi-dimensional formulation of the generic g order spline can be obtained simply by:

$$\varphi^{(g)}(\underline{t}) = \varphi^{(g)}(t_1, t_2) = \varphi^{(g)}(t_1) \cdot \varphi^{(g)}(t_2) \tag{6}$$

Figure 4 shows the behaviour of the first order bi-dimensional spline known also as bilinear spline.

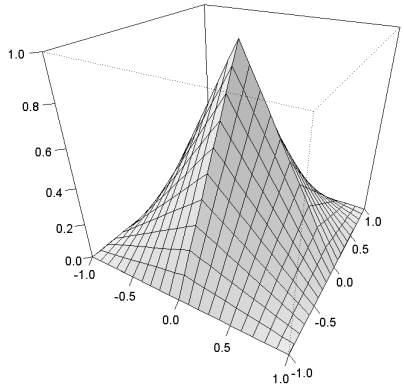


Figure 4. Bi-dimensional first order spline or bilinear spline

If we suppose that d($\underline{t}$) can be modelled as:

$$d(\underline{t}) = \sum_{k=1}^{N} \lambda_k \varphi^{(g)}(\underline{t} - \underline{t}_k) \tag{7}$$

the spline coefficients $\{\lambda_k\}$ can be estimated from the corresponding observation equations:

$$d_0(\underline{t}_i) = \sum_{k=1}^{N} \lambda_k \varphi^{(g)}(\underline{t}_i - \underline{t}_k) + v_k \tag{8}$$

by using the classic least square estimation method.
This ordinary spline interpolation approach suffers a rank deficiency problem when the spatial distribution of the data is not homogeneous. To make evident this concept, in figure 5.a a sample of 7 observations and the first order splines, whose coefficients we want to estimate, are shown. With this data configuration the third spline can not be determined because its coefficient never appear in the observation equations: the unacceptable interpolation results is shown in figure 5.b.
The simplest way to avoid this problem is to decrease the spline resolution with the consequent decreasing of the interpolation accuracy, especially where the original field d(t) shows high variability.
Since homologous points detected on geographical maps are usually not regularly distributed in space, the use of single resolution spline functions leads to two different scenarios.
In the first one, with low resolution spline functions, the interpolating surface is stiff also in zones where a great amount of points is available.

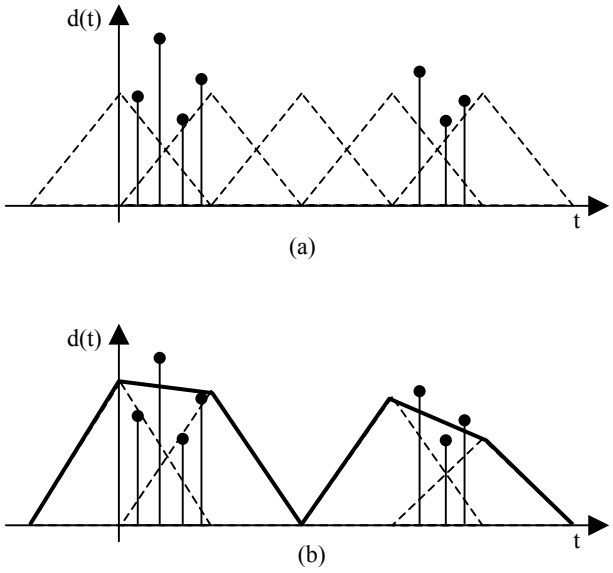


Figure 5. Examples of mono-resolution spline interpolation: data (a) and interpolation (b)

On the opposite, in the second case, corresponding to high resolution spline functions, a more adaptive surface is obtained but the lack of points in some area can give rise to local phenomena of rank deficiency, making the interpolation unfeasible. The multiresolution approach removes this problem.

2.2.2 The multi-resolution spline interpolation approach

The main idea is to combine splines with different domain dimension in order to guarantee in every region of the field a resolution adequate to the data density, that is to exploit all the available information implicitly stored in the sample data.
To show the advantage of this approach we suppose to interpolate the mono-dimensional data set shown in figure 6.a.
The classic spline interpolation approach requires to use a grid resolution in such a way that every spline coefficient appears at least in one observation equation. Figure 6.b shows the maximum resolution interpolation function which is consistent with the data set. The constraint on the grid resolution avoid the interpolation function to conform to the field data in high variability locations. Moreover, the use of the smallest allowable resolution can make the estimations sensible to the single observations in regions where data are sparse; the consequence is the generation of unrealistic oscillations, due to the fact that the noise is insufficiently filtered.

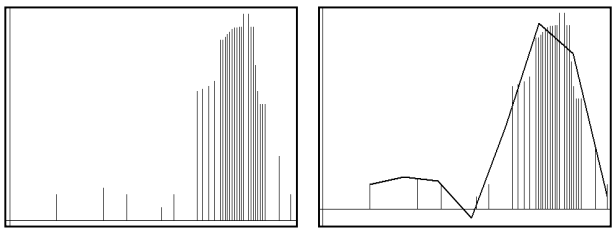


Figure 6. Sample data (a) and result of spline interpolation using mono-resolution approach (b)

In one dimension the multi-resolution can be obtained by modelling the interpolation function d(t) as:

$$d(t) = \sum_{h=0}^{M-1} \sum_{k=0}^{N_h-1} \lambda_{h,k} \cdot \varphi_{\Delta_h}(t - k\Delta_h) \qquad (9)$$

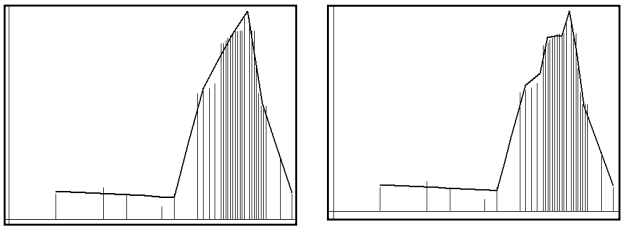
where $\varphi_{\Delta h}(t)$ is the first order spline with h resolution on the domain $[-\Delta_h, \Delta_h]$; M is the number of different resolutions used for the interpolation (levels); $\lambda_{h,k}$ is the k^{th} spline coefficient at resolution h; N_h is the number of spline with resolution h; Δ_h is the half-domain of the spline at the resolution h. In order to uniformly distribute the spline into the whole domain $D=[t_{min},t_{max}]$ the (9) can be rewrite as follow:

$$d(t) = \sum_{h=0}^{M-1} \sum_{k=0}^{N_h-1} \lambda_{h,k} \cdot \varphi\left(\frac{2^{h+1}(t - t_{\min})}{(t_{\max} - t_{\min})} - k \right) \qquad (10)$$

The model requires the imposition of constraints on the λ coefficients to guarantee the occurrence of a spline only in locations where data are enough to its coefficient estimation and to avoid the contemporaneous presence of two or more splines, obviously at different resolution, at the same grid position. The second event in fact causes the singularity of the normal matrix in the least square estimation. In order to avoid it, the i spline with resolution h_i is activated in point t_i, and therefore the $\lambda_{h,i}$ coefficient is not zero if:

- at least f observations are located inside its definition domain;
- no j spline exists with resolution h_j such as $t_i = t_j$ and $h_i<h_j$.

The f parameter acts as filtering factor to be used in the interpolation to avoid singularity.
The results of two multi-resolution spline interpolations are shown in figure 7.



(a) (b)

Figure 7. Results of multi-resolution spline interpolation with 4 (a) and 5 (b) levels

The bi-dimensional formulation can be directly obtained generalizing the mono-dimensional case.
We suppose that $d(\underline{t}) = d(t_1,t_2)$ can be modelled as:

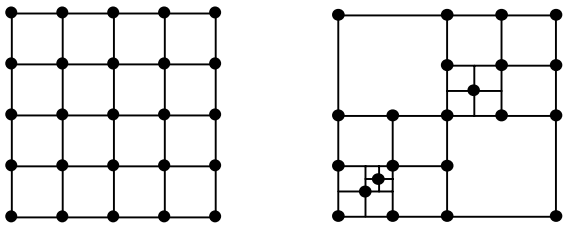
$$d(\underline{t}) = \sum_{h=0}^{M-1} \left[\sum_{l=0}^{N_{1h}-1} \sum_{k=0}^{N_{2h}-1} \lambda_{h,l,k} \; \varphi_{\underline{\Delta}_h}(\underline{t} - \underline{\Delta}_h \underline{\tau}_{lk}) \right] \qquad (11)$$

where:

$$\underline{\Delta}_h = \begin{bmatrix} \Delta_{1h} & 0 \\ 0 & \Delta_{2h} \end{bmatrix}$$

Δ_{1h} = x grid resolution;
Δ_{2h} = y grid resolution;
$\varphi_{\underline{\Delta}_h}(\underline{t}) = \varphi_{\Delta_{1h}}(t_1) \cdot \varphi_{\Delta_{2h}}(t_2)$
M = number of different resolutions used in the model;
$\underline{\tau}_{lk} = [l\ k]^T$ = node indexes (l,k) of the bi-dimensional grid;
$\lambda_{h,l,k}$ = coefficient of the spline at the grid node $\underline{\tau}_{lk}$;
N_{1h} = number of x grid nodes at the h resolution;
N_{2h} = number of y grid nodes at the h resolution.

As in the mono-dimensional model, a spline only starts up where data are enough to its coefficients estimation. Moreover, as usual, for each grid node only one spline is defined. It is important to notice that even though in a multi-resolution approach there is one grid for each level of resolution, the grid at resolution i+1 is built starting from the grid at resolution i by halving the grid step in both directions. The i-resolution grid nodes are therefore a subset of the ones at i+1 resolution (see figure 8).



(a) (b)

Figure 8. Interpolation grid with mono (a) and multi-resolution (b) approach

2.2.3 The automatic resolution choice

By passing from N to N+1 interpolation levels we introduce a certain number of splines whose coefficients, computed by least square estimation, are not null because of the stochastic deviation due to the noise. It is necessary to consider if the contribution of these new splines is significant, that is if they add new information to the field modelling or they only "chase" the noise. Given:

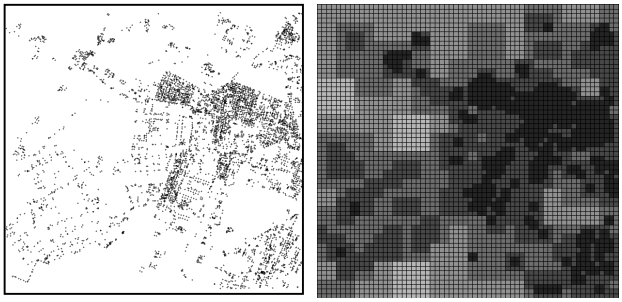
n_1 = number of spline used with N levels;
n_1+n_2 = number of spline used with N+1 levels;
S_{N+1} = set of coefficients of the new n_2 splines;

we find the N level such as the statistical hypothesis:
$H_0 : \{s_k = E\{\hat{s}_k\} = 0 \quad \forall s_k \in S_{N+1}\}$, being $E\{\cdot\}$ the expectation operator, is accepted. Without detailing the test, from the deterministic and stochastic model of the least squares approach we compute a variate F_0 which can be compared, with a fixed significance level α, with the critical value F_α of a Fisher distribution of $(n_2,N-(n_1+n_2))$ degrees of freedom. The test to accept the hypothesis H_0:{N is the resolution to choice, that is the increase of the spline number with n_2 new splines does not improve the fitting model and therefore their coefficients are null} can be formulated as follow: if H_0 is true then F_0 must be smaller than F_α with probability (1-α), otherwise H_0 is false and we have to iterate the test with N+1 resolution levels.

2.3 Application example

The procedure presented in the previous paragraphs was tested both on artificial scenarios appropriately designed for estimate

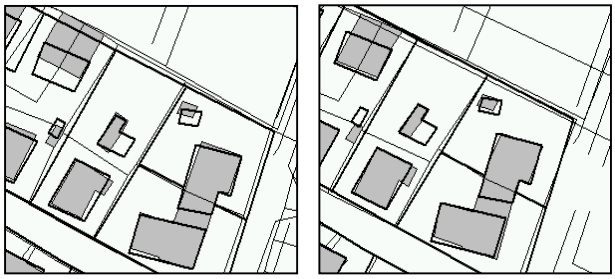
its performance and on real situations. The example here proposed is a real case: the maps that have to be combined into a single system are a cadastral one at scale 1:1000 and a regional one at scale 1:5000. An area of about 4 Km^2 is represented on the maps. The transformation was applied on the homologous points automatically detected by the procedure previously mentioned. In figure 9 the spatial distribution of the homologous points and the corresponding multi-resolution splines are shown. The five different resolutions are highlighted with different grey gradation (higher resolution are darker). It is important to notice that the heterogeneous distribution of the control points makes in this case inapplicable the mono resolution spline interpolation, at least if we are trying to locally model the differences between the two maps.



(a) (b)

Figure 9. Spatial homologous points distribution (a) and the multi-resolution spline collocation (b)

The results obtained by using a multi-resolution approach are, as expected, better than those due to the classic affine transformation. To have an example of the improved performances in figure 10 a detail of the overlaps between the two maps by using the classic affine transformation and the multi-resolution spline approach is shown.



(a) (b)

Figure 10. Overlay of two maps using the affine transformation (a) and the multi-resolution spline approach (b)

It is evident that the localized deformation in the upper-left corner of the map has been "catched" by the multi-resolution transformation (b) with the consequence of the improved overlap between the two maps.

3. CONCLUSIONS

The use of spline functions in modelling deformations between maps, compared to affine or polynomial interpolation, allows to have a greater number of coefficients to make more adaptive and localized the transformation. The multi-resolution approach here presented removes the rank deficiency problem that ordinary spline approach suffers for. Moreover a statistical test allows to choose the level of multi-resolution to be adopted in order to better model the deformations between the two maps.

REFERENCES

Cobb, M.A., Chung, M.J., Foley, H., Shaw, K., Miller, V., 1998. A rule based approach for the conflation of attributed vector data. *Geoinformatica*, 2(1), pp. 7-35.

Doytsher, Y., Gelbman, E., 1995. A rubber sheeting algorithm for cadastral maps. *Journal of Surveying Engineering – ASCE*, 121(4), pp.155-162

Doytsher, Y., Hall, J.K., 1997. Gridded affine transformation and rubber sheeting algorithm with Fortran program for calibrating scanned hydrographic survey maps. *Computers & Geosciences*, 23(7), pp. 785-791.

Doytsher, Y., 2000. A rubber sheeting algorithm for non-rectangular maps. *Computer & Geosciences*, 26(9), pp. 1001-1010.

Gabay, Y., Doytsher, Y., 1994. Automatic adjustment of line maps. Proceedings of the GIS/LIS'94 Annual Convention 1, pp. 333-341.

Gabay, Y., Doytsher, Y., 1995. Automatic feature correction in merging of line maps. Proceedings of the 1995 ACSM-ASPRS Annual Convention 2, pp. 404-410.

Gillman, D., 1985. Triangulations for rubber sheeting. In: Proceedings of 7th International Symposium on Computer Assisted Cartography (AutoCarto 7), pp. 191-199.

Laurini, R., 1998. Spatial multi-database topological continuity and indexing: a step towards seamless GIS data interoperability. *International Geographical Information Science*, 12(4), pp. 373-402.

Lupien, A., Moreland, W., 1987. A general approach to map conflation. In: Proceedings of 8th International Symposium on Computer Assisted Cartography (AutoCarto 8), pp. 630-639.

Moritz, H., Sunkel, H., 1978. Ap*proximation Methods in geodesy*, Herbert Wichmann Verlag, Karlsruhe.

Rosen, B., Saalfeld, A., 1985. Match criteria for automatic alignment. In: Proceedings of the 7th International Symposium on Computer Assisted Cartography (AutoCarto 7), pp. 456-462.

Saalfeld, A., 1985. A fast rubber-sheeting transformation using simplicial coordinates. *The American Cartographer*, 12(2), pp. 169-173.

Saalfeld, A., 1987. Triangulated data structures for map merging and other applications in geographic information systems. In: Proceedings of the International Geographic Information Symposium, III, pp. 3-13.

Saalfeld, A., 1988. Conflation - automated map compilation. *International Journal of Geographical Information Systems,* 2(3), pp. 217-228.

White, M.S., Griffin, P., 1985. Piecewise linear rubber-sheet map transformations. *The American Cartographer,* 12(2), pp. 123-131.

Yaser, B., 1998, Overcoming the semantic and other barriers to GIS interoperability. *International Geographical Information Science*, 12(4), pp. 299-314.

Yuan, S., Tao, C., 1999, Development of conflation components, Geoinformatics and Socioinformatics, The proceedings of Geoinformatics '99 Conference. http://www.umich.edu/~iinet/chinadata/geoim99/Proceedings/yuan_shuxin.pdf (accessed 27 April 2004).