# Accuracy Improvement for Italian Digital Cadastral Maps

Alberto Cina*
Ambrogio Manzino*
Tamara Bellone*

* DIATI, Politecnico di Torino

*The original implant map (OI) of Italian Cadaster was first made out of topographical measurements.*
*These maps have never been used for public consultation, but they are the most accurate ones. The publicly accessible maps are named "visure maps" (VM) and were obtained from the original implant maps and consistently updated. However, these maps have suffered many deformations due to continuous use over time.*
*These VM were then vectorialized at the beginning of the 2000s and have therefore undergone a further process of degradation. The goal of this work is to improve the accuracy of the Italian VM using the metric informations. This can be achieved if we can precisely model the deformations, measurable by the recognition of boundaries and details still present on both maps and, after this, to apply this inverse model to the visure maps.*
*The illustrated procedure is adaptable to other cartographies with similar problems.*
**Keywords:** *Cadastral Maps, digital matching, Deformation models, Kriging.*

***Il miglioramento delle accuratezze delle mappe vettoriali del catasto.*** *La mappa originale d'impianto (OI) del catasto italiano fu inizialmente realizzata con misure topografiche. Queste mappe non sono mai state utilizzate per la consultazione pubblica, ma sono le più accurate.*
*Le mappe accessibili pubblicamente sono denominate "Copioni di Visura" (CV) e sono state ottenute dalle mappe originali d'impianto e costantemente aggiornate nel tempo. Tuttavia, queste mappe hanno subito molte deformazioni a causa dell'uso continuo nel tempo. I CV sono stati vettorializzati all'inizio degli anni 2000 e hanno quindi subito un ulteriore processo di degrado.*
*L'obiettivo di questo lavoro è migliorare l'accuratezza di queste mappe catastali CV utilizzando le informazioni metriche ottenute dagli OI . Questo si ottiene modellando con precisione le deformazioni, misurabili mediante il riconoscimento dei confini e dei dettagli ancora presenti su entrambe le mappe e, successivamente, applicare il modello deformatico inverso alle mappe visura.*
*La procedura illustrata è adattabile ad altre cartografie con problemi analoghi.*
***Parole Chiave:*** *catasto, mappe originali d'impianto, correlazione, modelli deformativi, kriging.*

## 1. Introduction

At the outset we must provide both some historical background and some specific terminology little known outside the sphere of cadastral survey, that is to say the application of cadastral maps to redefine boundaries to subdivide plots of land and other operations.

The Italian Cadastre, at present incorporated, somewhat short-sightedly into the National Revenue Service, has a huge collection of large and very large scale maps (from 1:4000 to 1:500, generally speaking) which is the fruit of almost a century of direct land surveying. (www.agenziadelterritorio.it). The introduction of the Italian Cadastre in 1865 made it possible to produce not only many geodetic and property value studies but also the trigonometric cadastral network, the original cadastral maps, the Registry of Land Parcels and the Buildings Registry together with their respective valuation estimates (Paroli, 1948). This great endeavour laid the foundations for metrical and economic management of real property. At the end of the last century, digitization of cadastral data began, including vectorization of those maps available for public viewing in local cadastral offices. Obviously, these maps were generally very different from the original ones due to various changes, including subdivisions or incorporations made in the course of time.

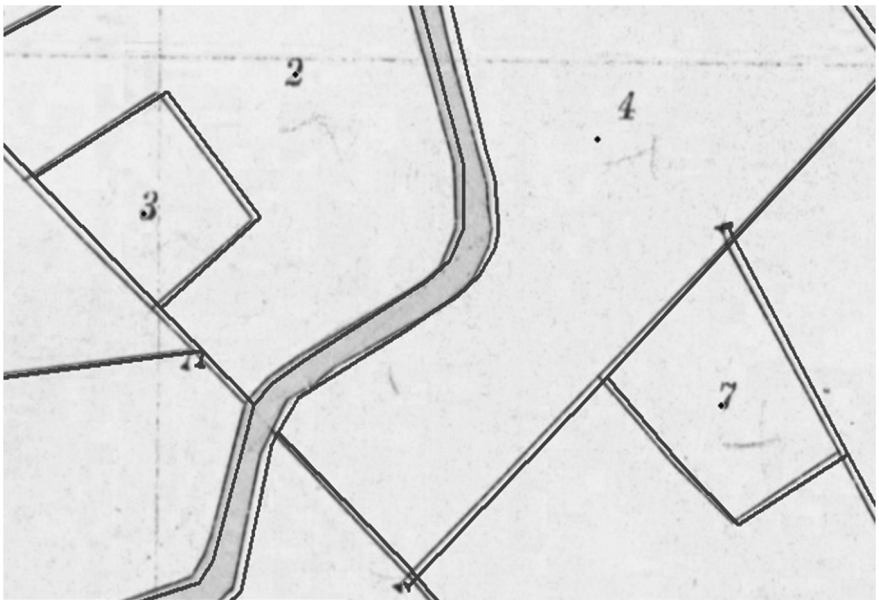


Fig. 1. Distortion of the vector map appears in green over the original beneath.
*Distorsione della mappa vettoriale in verde rispetto all'originale d'impianto sottostante.*

The physical paper support of these reference maps derived in many cases from a contact copy on "*carta canapina*" (paper reinforced with hemp fibres) from a zincographic plate in turn obtained from the original map.

It is self-evident that, in the course of the successive procedures which have contributed to the final production of the digital cadastral map, many distortions have occurred due to the numerous operations which can be summarized thus:

- From the measurements to the creation of the original map.
- The creation of the cadastral reference map plates.
- Additions, updatings and deformations on the physical reference maps.
- Outliers, both systematic and accidental in the acquisition and manual digitization of cadastral reference maps.

First run deformation recovery today makes little sense and would, in any case, be almost impossible: rarely, and in cases of serious disputes, does one refer to the "original relief sketches", today almost impossible to locate. We assume as our most reliable metrical basis the original map. Even this, drawn on a very robust parameterized paper support, has been subject to slight deformations over time. Recently these maps have been digitalized with high-precision flat scanners, ultimately yielding raster maps which still display small distortions, but these, thanks to the grid present on the map (fig. 1), may be easily recovered.

Discussion has been made on how it is possible to recover such distortions with an accuracy similar to graphical error and with an even better precision (Cina *et al.*, 2010). We should highlight, in this regard, that only the original maps possess a complete parameterization over their entire surface areas for the purposes of deformation recovery. Henceforth, when referring to the original map, we will always mean the raster product obtained after the warping procedures which have recovered the slight deformations by means of automatic measurement of the entire parameter grid.

In this study we, therefore, focus on the recovery of those imprecisions due to steps 2, 3 and 4, assuming that the original cadastral map, once digitized and after recovery of the aforementioned distortions, is our precision metrical reference.

Some technicians operating in the cadastral field recommend, for optimal recovery of the metrical information on the vector maps of the cadastre, usually provided to users in CXF format, a direct transformation on the ground of points also identifiable on the map, usually referred to as "double points" to rototranslate the cartographic system into the system being used by the surveyor.

A similar procedure might, in our opinion, be adequate for very limited areas of a map but, in any case, it is both logically and practically weak, especially for more extensive areas.

What is observed might be very different from what it was necessary to include on the creation of the map, both in a topological sense, and because of extensions or modifications due to agreements among neighbours or usage tacitly producing a right. What may be adapted locally is ultimately different from what has occured in a general context.

It may make sense, if anything, in restoring boundaries, to refer to the written and/or topological information available. For example, a boundary which is defined by a mountain peak, the intersection of a canal and a road neither of which have ever been changed, information regarding a fence running parallel to the likely property boundary (Cina *et al.*, 2016).

Returning to our proposal, a feature which is measured on both maps, if recognizable, is certainly a representation of the same feature on two different supports, thus no topological error is committed. Obviously, many features present on the vector map will not be present on the most ancient original map. More rarely, due to demolitions or incorporations, features present on the original map do not appear on the digital vector map.

## *2. Method*

Before considering the method for partial recovery of the distortions, it is necessary to measure the features in the most precise and extensive way possible. The simplest approach is to insert both maps in a GIS or CAD system capable of handling raster and vector maps together and, for every point recognizable by the naked eye as being unmodified, to measure the coordinates on both the first and second level. The advantage of such a procedure is that it is human intelligence which decides whether a pair of points (generally the edge of the said feature) is actually "homologous". However, if we wish to measure all those points which are recognizable on both levels, the procedure may prove extremely lengthy and, in any case, highly dependent on the expertise of the individual operator. For our purposes, the lowest error of measurement possible on the raster map is negligible: even with an acquisition at 300 d.p.i., pixel size (0.08 mm) is inferior to graphical error. We realize, however, that, with this procedure, what must be measured is always constituted by intersections, for example, the

boundaries of a parcel of land, building edges (beware extensions!), the intersections of different features.The challenge is therefore to recognize automatically these edges both at vector and raster level (Gonzales and Wodd, 2002). Often, no similarities will be observed. While the procedure does "designate" the points on the two maps as homologous, it will require, nonetheless, robust procedures to exclude those similarities which are metrically incompatible or meaningless.

### *2.1. Homologous or pseudo-homologous points?*

The basic idea behind automatic recognition of homologous points at raster and vector levels is simply that of also transforming the CXF digital map into a raster image so that the matching algorithms may be applied at pixel or subpixel level on the digital photogrammetry. For the sake of correctness, since the two raster images produced are not of the same object, it is more appropriate to define edges and features on both maps as "pseudo-homologous". Hereafter, we will often use the adjecive "homologous" with this meaning for which we ask the reader's forgiveness.

The search algorithms for homologous points on two raster maps may be found among those on the digital photogrammetry (Kraus, 2011): we asked ourselves, therefore, whether, also in this particular instance of "cartographic" images, these methods may be utilized.

In the first analysis two algorithms were evaluated: "Harris Corner Detection" and "SURF". Harris Corner Detection (HCD) (Harris and Stephens, 1988) is based on a function of local autocorrelation $c(x, y)$ (1). which by shifting a patch measures signal changes of small quantities ($\Delta x$, $\Delta y$):

$$c(x,y)=\sum_W \left[ \begin{matrix} I(x_i,y_i)+ \\ -I(x_i+\Delta x, y_i+\Delta y) \end{matrix} \right]^2 \quad (1)$$

where $I(x_i, y_i)$ is the image function and $(x_i, y_i)$ are the coordinates of the point in window $W$ centred in $(x, y)$.

SURF (*Speeded Up Robust Feature*) (Bay *et al.*, 2006) is a robust detector of image characteristics, based on wavelet technologies. In part, it takes its inspiration from the SIFT, Scale-invariant feature transform algorithm (Lowe, 1999), but is quicker.

Once the points of interest on the two images have been defined, the correspondences are to be sought with robust methods capable of considering the *inliers* alone. The best-known technique of outlier elimination is that belonging to those metodologies defined RANSAC (*RANdom SAmple Consensus*) (Fishler and Bolles, 1981) which use efficient iterative methods even where outliers are numerous.

Given the characteristic nature of cadastral maps, which might, in some cases, have houses or parcels with repetitive *patterns*, experimentation has been performed using the HCD, SURF and RANSAC techniques provided by the Matlab® tool even on two photograms of a rather flat facade with many wall tiles.

The search operation to find the homologous points is performed correctly (fig. 2) and it is possible to eliminate the deformations, in this case ones of prospective, on the basis of the numerous double points found.

We were curious to know, at this point, whether analogous results could also be obtained automatically between the two raster images of the two maps. Unfortunately, the answer is no. The search for homologous points provides several thousand candidates on different maps but, alas, subsequent to *matching*, the RANSAC technique produces an empty set!

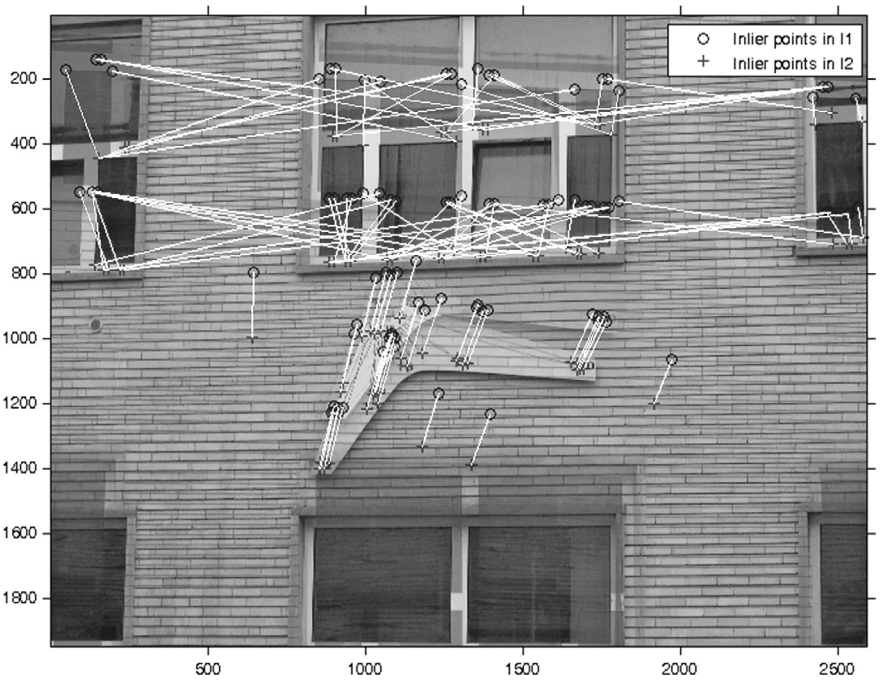


Fig. 2. Right and left photograms superimposed with complementary colour: inliers and parallaxes.
*Fotogrammi destro e sinistro sovrapposti a colori complementari: inliers e parallassi.*

Such maps do not produce pairs of points which may be considered homologous.

The explanation is visible in Fig. 3, where it is clear that the numerous points found are not the right ones for a transformation, even if they are close to the points of interest. In general, they are written elements, and other graphic symbols, impossible to eliminate from the digital cadastral maps.

Analogous results, for the sake of brevity not displayed here, were obtained by applying the SURF algorithm.

The path taken is then that of seeking maximum correlation among the points comprised within search windows on the two raster maps on the basis of the coordinates of the edges extracted from the vector file. We chose to use the area around the coordinates extracted by the vector file as reference image of dimensions (s, s) and the corresponding areas on the original map as search image of dimensions (r, r). From the analyses conducted it was deemed reasonable to adopt variable pixel sizes for the reference from a minimum of (9, 9) to (13, 13). The size of the search image matrix is roughly double (fig. 4).

The search for homologous points starts out from the coordinates read in the CXF vector file and converted into pixel coordinates (i, j) through the world file (TFW, or JGW) both on the rasterized vector map, and on the original map (fig. 4).

The points found are very numerous but, unfortunately, many are not significant for the purposes of the transformation, being written elements or blotches which happen to appear in the search interval (fig. 5). Some of these errors are, however, easily recognized when minimum thresholds of the linear correlation coefficient are set.

It is important then to establish a minimum acceptance thre-

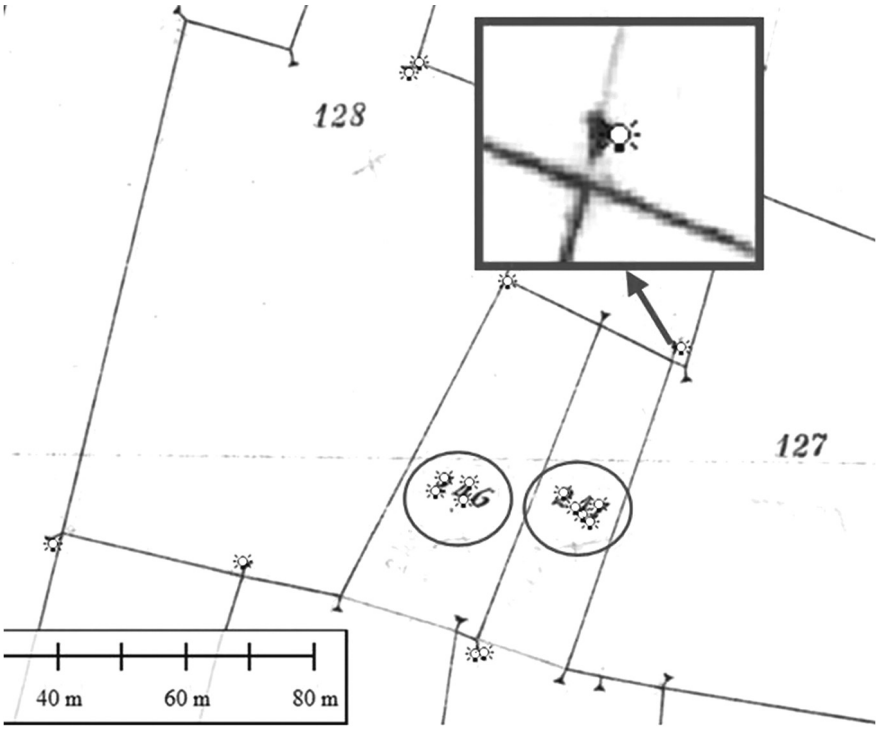


Fig. 3. Some points found by HCD (lightbulb) and rejected by RANSAC.
*Alcuni punti ritrovati da HCD (lampada) e scartati da RANSAC.*

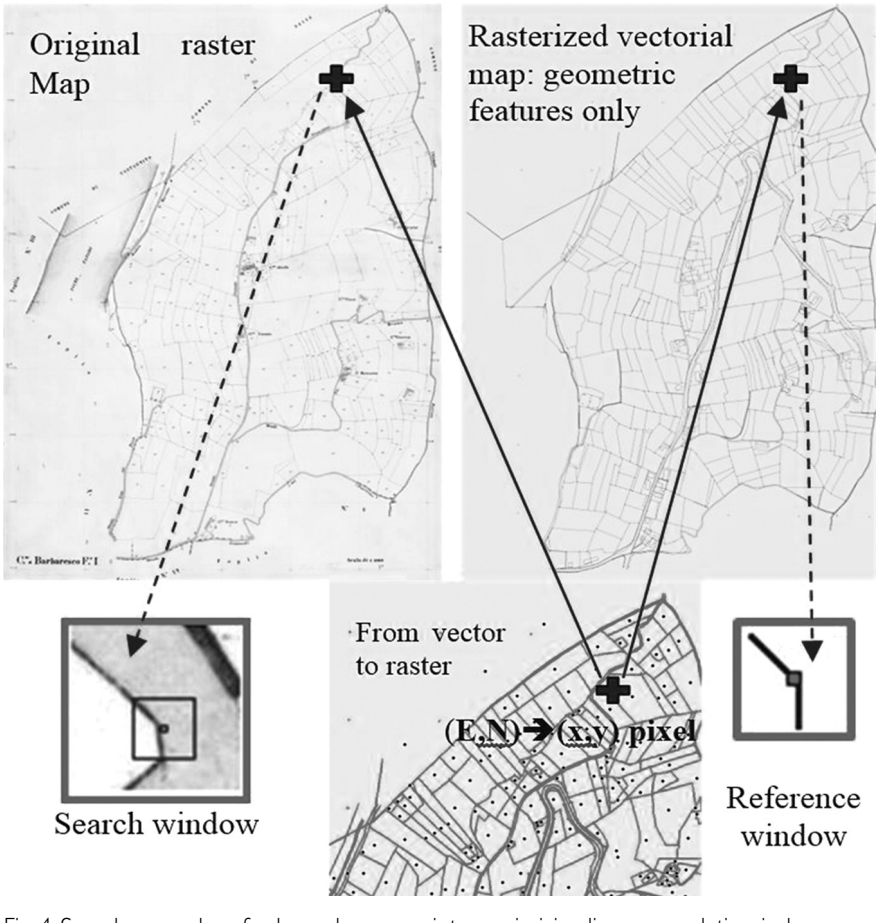


Fig. 4. Search procedure for homologous points maximizing linear correlation index.
*Procedimento di ricerca di punti omologhi che massimizza l'indice di correlazione lineare.*

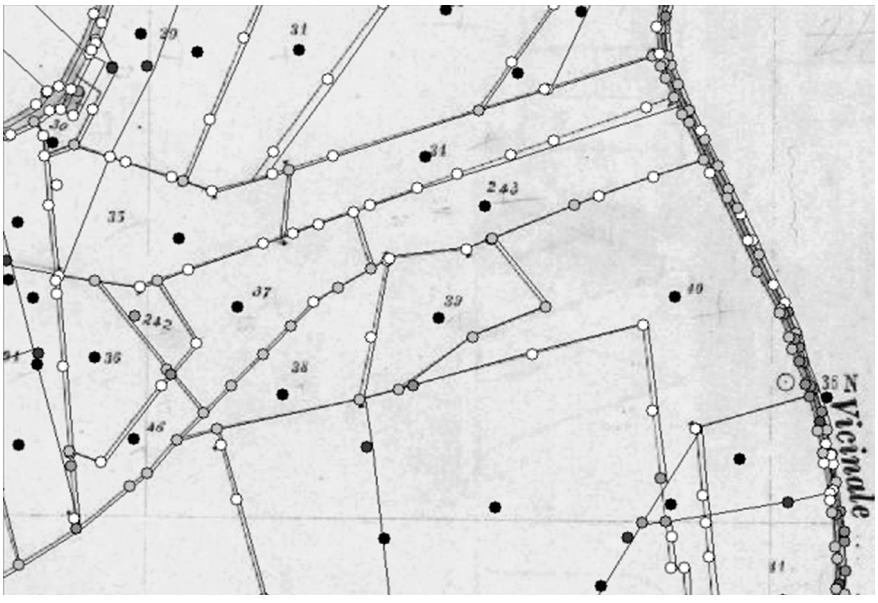


Fig. 5. Homologous points found and linear correlation index ρ: in black ρ < 0.05, in gray ρ > 0.36, in white intermediate values.
*Punti omologhi ritrovati e indice di correlazione lineare r: in nero ρ < 0.05, in gray ρ > 0.36, in bianco valori intermedi.*

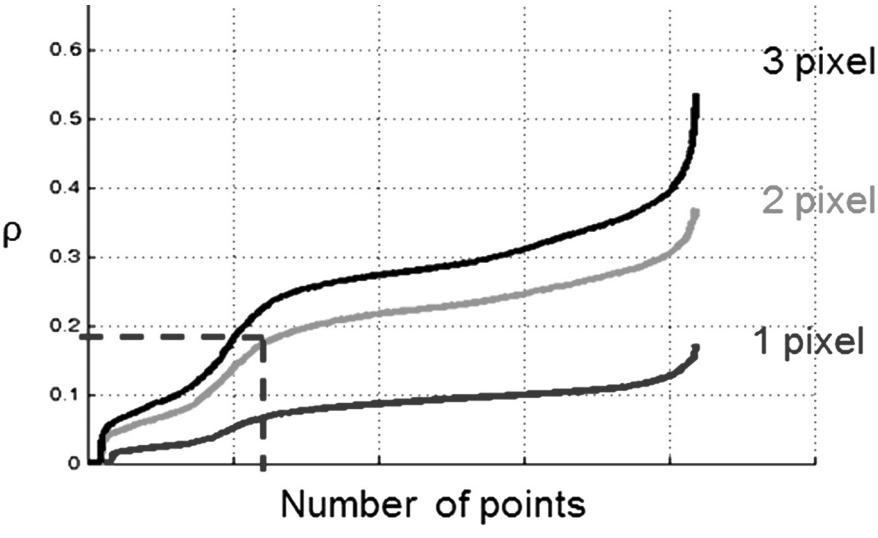


Fig. 6. Linear correlation index with variation in line width.
*Matrici sagoma in toni di grigio (a), in bianco e nero (b) e matrice ricerca (c).*

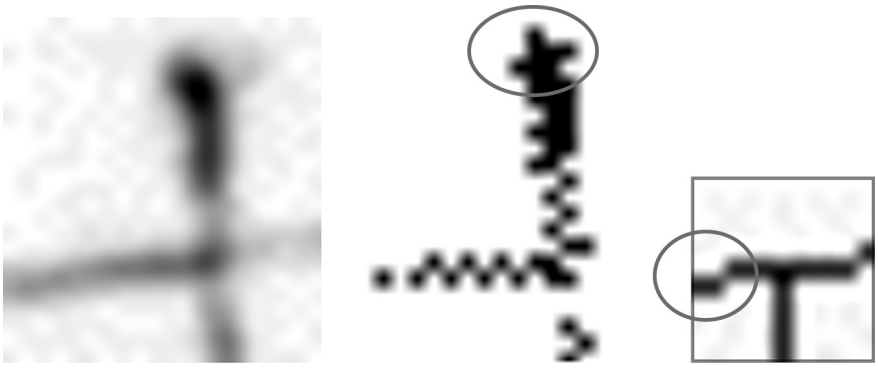

Fig. 7. Reference matrices in grey scale (a), in black and white (b) and search matrix (c).
*Indice di correlazione lineare al variare dello spessore della linea.*

shold: this, however, depends on the pixel size adopted in the process of rasterizing the vector map, shows the number of points excluded from the recognition when modifying the minimum value of the linear correlation coefficient where rasterization is performed with lines of a width of 1, 2, and 3 pixel aty 300 dpi. For example, on about 22,000 points present in the vector file, if one seeks to limit to 700 the number of points rejected, one may choose a value of ρ = 0.18 with a line of 2 pixels in width.

Moreover, the width of the lines may not be greater than the graphical error, conventionally taken to be 0.2 mm, even using subpixel correlations and, as a lower limit, itcannotbe very different fromthe average width of the lines on the map for comparison. If it were thus, if we were to adopt, for example aa width of 1 pixel, not only would a good percentage of candidates be rejected, but the best matching would be found in the darkest parts of the search without corresponding globally to the form sought.

The three images in fig. 7 show in enlargement the search matrix on the left and reference matrix on the right. Should we adopt the width of a single pixel for the reference window, the best matching occurs in the upper part of the search matrix: this is little evident in the grey-shaded image, but is more obvious if the search image is transformed into B/W, as in Fig. 7b, reduced to two colours just for the purpose.

The search for the limit value of the correlation index does not, however, fully exhaust the search for outliers: some resist this first selection and this requires the issue of the transformation to be handled with robust estimation algorithms, based on the concept that, on average, between the two maps there may be a global defor-

mation, largely a translation, such as seen in Fig. 1 and which may be recovered better through a six-parameter affine transformation, (2) added to a second deformation which may be modelled stochastically.

$$\begin{pmatrix} X_v \\ Y_v \end{pmatrix} = \begin{pmatrix} X_0 \\ Y_0 \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} X_i \\ N_i \end{pmatrix} \qquad (2)$$

These parameters are to be calculated robustly and applied for a more refined search both of the homologous points and the outliers (Doytsher and Hall, 1997).

The previous experiments formed the basis of the complex procedure, which we will now describe, which makes possible the automatic recognition on the two maps of homologous points, provided that a series of preliminary conditions are met, which are:

- The resolution of the rasterized vector map must be the same as the original map, with the same system of reference and with the same limits. In the trials performed, prove resolutions of 20 cm and 40 cm were tested, these being equivalent on the ground for both maps on scale 1:2000.
- The starting coordinates and those for comparison of the vector map are not those of the rasterized image but the numerical/digital ones extracted from the CXF file which have undergone the unknown deformations to shape.
- Prior to rasterization it is necessary to strip the CXF file of all symbols, of toponyms and of any other detail which is not geometric but merely descriptive, eliminating colours used to describe features.
- The size, the width of the lines in the rasterized image must be generally equal to the width of the lines from the image of the original map.
- The procedure must perform a series of controls which concern not only the threshold values of the matching index, but also the existence in the two matrices of edges or significant intersections;
- The procedure, considering a global deformation which may be modelled with (2). after having eliminated outliers, may be repeated with reduced sizes of reference and search windows and with the use of transformation parameters obtained from the previous step.

## *2.2. Optical and geometric problems in the two images*

As regards the vector map, from the tests performed it was concluded that the main focus for rasterization is that of ensuring that the average width, in pixels, of the synthetic line be similar to that on the raster map. In this operation, one may, however, choose to produce an image in greyscale or in black and white (B/W). The number of homologous points found and the accuracies obtained in both cases are similar. It was preferred, nonetheless, to use greyscale images with 8 bit depth as this reduces the "blocking" effect due to resampling, since darker-shaded pixels are closer to the ideal line of the vector source.

A more refined treatment is, however, reserved for the original map, or rather, for the search windows extracted from that map. The starting image is in colour on a very time-yellowed background (fig. 8). Nor do the lines always have a sharp contrast against the background. In addition, line width is uneven. To transform it into greyscale using the three colours is the least opportune choice; what was done, however, was to highlight the information contained with the method of the principal components. The only contraindication, in rare cases only, is the proximity to the line of a sharp element written in another colour, as, for example, the name of a street or a symbol. This too is highlighted and, on the resulting image, the correlation index on such symbols or written elements may be extreme, albeit without any geometric significance. As mentioned already, experimentation was performed, both for the reference matrix and for the sear-

Fig. 8. Two portions of the same map, the original cadastral one (a) and the vector one (b) are in appearance profoundly different in geometry and radiometry.
*Due porzioni della stessa mappa, quella di impianto (a) e quella vettoriale (b) sono in apparenza profondamente diverse per geometria e radiometria.*

ch matrix (on the original map) to transform the images into B/W, that is to say, 2 bit, after appropriate calibration of the cut thresholds. The results were no better and, as such, the correlations were always made with images in greyscale with 8 bit optical depth.

### *2.3. Dissimilarities*

The dissimilarities between the two images which represent the original map and the rasterized vector map, which are reflected in the reference and search matrices, arise for the reasons for which we have called such images "pseudo" homologous and which we outline:

1. A new feature in the reference, derived from the vector is not present in the search (subdivisions);
2. A new feature in the search: the original map, does not exist in the reference (incorporations);
3. Errors of digitization present in the CXF file.

In this third category we do not wish to include accidental errors, but rather those associated with the digitization of symbols with incorrect metrical value. In the great majority of cases, these are property boundaries.

This symbol, very similar to the letter "Y", (save for municipal boundaries), indicates at its base the presence of a boundary stone on the ground (fig. 10). This boundary does not always coincide with an intersecting of borders, more often it is set back from one of the two boundary lines.

On many occasions, this symbol on the vector map has been digitalized by translating it along the edge of the boundary. We regard this as being an error due to an inadequate understanding of the full metrical significance, as well as its symbolic connotation, on the part of the person who has vectorized these cadastral reference maps. (enlargement of Fig. 19). In order to avoid this type of error we preferred to exclude all the boundary features on the vector map before rasterization (fig. 9).

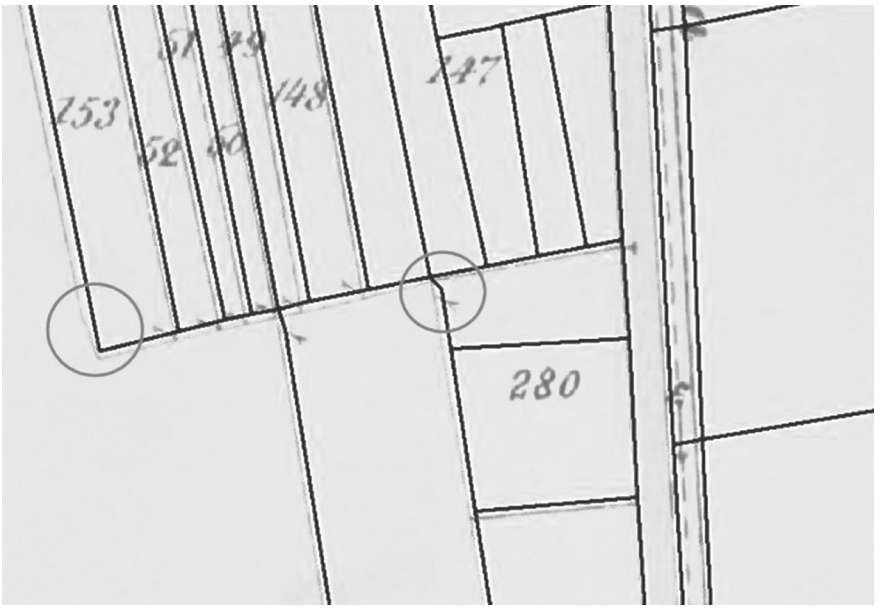


Fig. 9. The reduced vector superimposed on the original map: we notice new subdivisions for parcels 147 and 280 but also errors of digitization for parcels 153 and 280.
*Il vettoriale ridotto sovraimposto alla mappa di impianto: notiamo nuovi frazionamenti per le particelle 147 e 280 ma anche errori di digitazione per le particelle 153 e 280.*

Given that in case 3) outliers can only be managed using a robust procedure (Rousseeuw, 1987) the question is now, what happens in the software in cases 1), 2) during the course of matching.

In general, in the first two cases the objective is to correlate an edge, that is something with an acute angle acuto on an image, or a node on which several lines conconverge, with a straight segment which does not riporta the new geometric elements, onto another image.

The result is that matching is very successful, but not at one

Fig. 10. Boundary stone found and boundary stone unearthed (Cina *et al.*, 2016).
*Termine lapideo ritrovato e termine lapideo fuori terra.*

point, but ratther at a place of points, that is to say, on the whole straight segment. Obviously, a maximum, and a very high one at that, will be reached, both in the whole pixel procedure and subpixel one. The position of this maximum, nonetheless, is random and has nothing to do with the reference in the case 1) or with the search in the case 2).

Since these cases are not rare, it is necessary to recognize them and manage them.

There are the extreme cases, didactically interesting, which indicate a path to a solution. When the reference matrix represents a vertical or horizontal segment, it is easy to remark that this matrix itself, or the design matrix which represents the greyscale and their derivatives do not completely fill the columns.

The same thing happens, strictly speaking, also for the reference matrix.

When the segment is inclined but edgeless, this problem is evidenced by the very high norm of these matrices, or the norm of the normal matrix (Strang, 2016). This information has then been used as a filter, after appropriate calibration, in order to exclude these points from the comparison a priori.

More frequent is the case illustrated in Fig. 9, when there exist, that is, new features not present on the original map, such a in parcel 147. In pratice, it is almost never possible to resolve the problem by searching for a filter linked to the norm of the search matrix. The image in the search matrix does not have well-defined lines such as for the reference matrix and it is rarer that the test is not passed.

Considerable time has ben spent trying to settle the issue with the criterion connected to the Harris number, which uses together the determinant of the matrice and its norm, but even in this case calibration proves almost impossible. Lastly, we used the search for the existence of the edge with maximum significance within the reference matrix, once again following the Harris criterion. In this case, too, n adequate calibration of the parameter of significance is required, but, ultimately, the results were acceptable.

A common feature of cases 1) and 2) is that linked to the progression of the linear correlation coefficient, which remains good, but ordering it from minimums to maximums, we can see that there are no peaks, but rather almost constant values close to the maximums. For this reason, a filter, linked to the minimum slope of the linear correlation index after ordering was applied.

It remains for us to mention the filter connected to the low correlation coefficient, $\rho min$.

Below this threshold the point is rejected. Lastly, in a percentage of no more than a few thousandths, having passed all the filters and having achieved subpixel matching,it may happen that one of the two translations be greater or smaller than a pixel. In this situation, the position calculated with the whole pixel procedure is assumed to be correct.

The software developed indicates how many points it uses at the start and how many points it rejects for each of the reasons listed

It then builds a DXF file in which the points used are represented with different symbols, and the points rejected with a distinct digital code for each type of residual. This file can be used for superimposition onto the original map and also the vector map to get an idea about the problems arising from the comparison of the two maps and what these might depend upon.

## *3. Results*

It might seem strange that in applying the methods described, translated into three Matlab® softwares, we should speak generically of results. These depend on the regulation of the parameters of all the aforementioned filters, furthernore they may be statistically of little significance because the experiments have been performed on a single map. If, on the one hand, the parameters "of syntony" are multiple and give us hope that the procedure may be adapted to other cases, certainly this will not be able to happen automatically, requiring thus a specific calibration, at least for similar types of maps. We must still define what we mean by result. The first software produces three archives, the first containing six parameters of a general affine transformation, which exists between the two maps. A second archive is graphic and in the DXF format, and shows the placement of all points of the digital map in the new position obtained from the "robust" affine transformation, as well as, marked with a different symbol, the position of maximum matching of pseudo homologous points. The robust estimate procedure was that of least squares reweighted with the inverse of the residuals. A final archive contains, for the pseudo homologous points only, four columns: two are the starting coordinates (x, y) of the digital map, on the same line are indicated the distances in x and y between the digital coordinates obtained after the affine transformation and the original map coordinates of the point considered homologous, in graphic form in Fig. 11.

We will call these quantities "residual deformations" or "deformations". In other words, these last two columns are the residuals of the robust affine transformation


Fig. 11. The DXF file produced by the first software package identifies with a red circle the acceptable homologous points which are indicated with a recognition code and with a progressive number.
*Il file DXF prodotto dal primo software individua con un cerchio rosso i punti omologhi accettabili che sono indicati con un codice di riconoscimento e con numero progressivo.*

of the pseudo-homologous points on the two maps.

The data in this third archive are read by a second program which calculates the respective empirical variograms separatately for the deformations in x and y. At this point, the theoretical variograms are estimated and the Kriging estimate is applied separately for x and y.

Finally, a third software, for each point of the digital map, sums the deformations obtained from the affine transformations and the Kriging estimates of the residual deformations, thus constructing a digital map in the DXF format which adapts very well to the original cadastral map (fig. 19). We will now proceed to describe the results of these three stages.

### *3.1. Recognition and measurement procedures*

Two maps were used, having, post rasterization, the same borders and identical pixel size: Sheet 24 of the Municipality of Mortara on a scale of 1:2000 with 20 cm pixel on the ground. The first software known as PORI: Pseudo Omologo Riconoscimento (Pseudo Homologous Recognition), on the second "run", was used with the following parameters: Reference and search windows respectively of 11 and 19 pixels, minimum permitted matching $\rho_{min}$ = 0.65, minimum slope of correlation increase = 0.15, minimum weight in the robust least squares procedure = 0.85, maximum permitted condition number $cn_{max}$ = $1.5 \times 10^5$. The pseudo homologous points to correlate total 2169. With these parameters the results on the map in question are:

After the search and before affine transformation the pseudo homologous points found number 571. After robust affine transformation these are reduced to 449. The value of $\sigma_0$ of the procedure is 0.13m. The rms of the residuals of the affine transformation, also comprising the residuals on the points rejected are: rms(x) = 0.35 m; rms(y) = 0.37 m.

It is to be noted that in the general "robust"affine transformation, a point is completely rejected if the residual in x or y has a lower weight than the set threshold. After stabilization of weights, new parameters are recalculated using the classic least squares procedure on the points remaining only, all having the same weight. The rms of the residuals are calculated on all points, both those used to recalculate the six parameters and those points rejected as being underweighted.

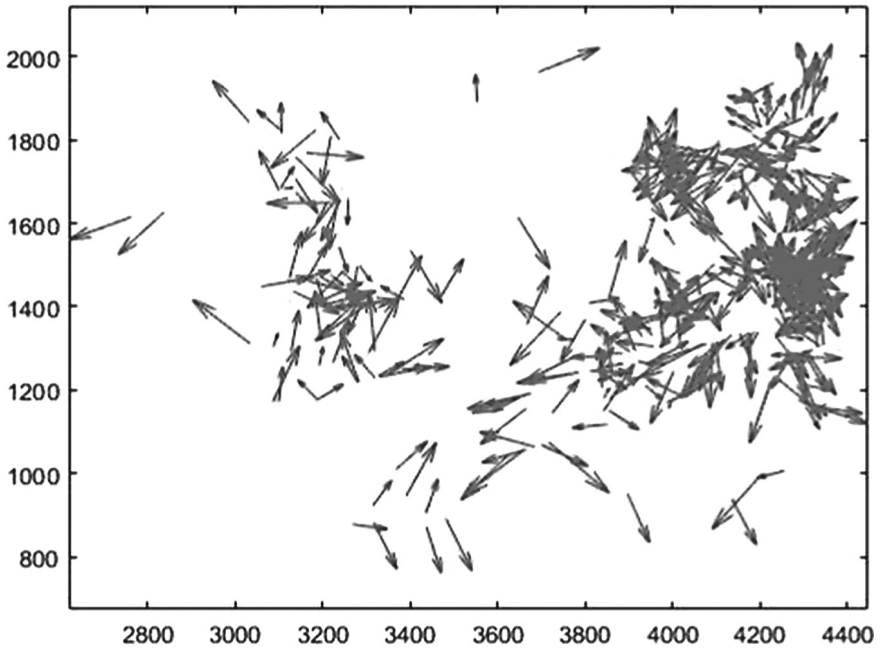


Fig. 12. Deformations amplified 200k times.
*Deformazioni amplificate 200k volte.*

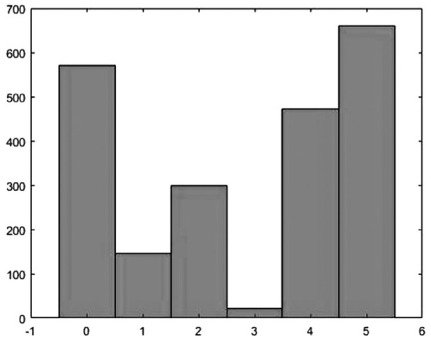


Fig. 13. Distribution of the exclusion categories of the points.
*Distribuzione delle categorie di esclusione dei punti.*

Figure 13, on the other hand, shows the classification of the 2169 points based on the exclusion criterion of the numerous filters. At level zero we find those 571 points never excluded and deemed homologous, added to which about 140 points at level 1 with a correlation coefficient below the set limit, and around 300 points with an increase correlation slope under 0.15. There are twenty or so points with "sub pixel" shifts in x or y of slightly over one unit. There are almost 500 points without significant edges within the search and a little under 700 points with a low condition number in the reference matrix.

### *3.2. Kriging estimation of deformations*

As shown by the previous results the deformations appear to be distributed unevenly, becoming denser in the urban centres but also apparently in a random fashion. In reality, this affirmation must be verified by means of an auto-correlogram or of a variogram (Cressie, 1993)].

The choice of using a stochastic method to estimate these residuals derives from a series of advantages (Oliver and Webster, 2015), which in our opinion are:

– It may be understood whether and to what degree these residuals may be white noise or in some way explained;
– From a statistical point of view Kriging is the best possible linear estimate
– Even if the variogram model function is a rough approximation of the phenomenon, the estimate with the use of such functions protects and filters the peak values where outliers are present.

For this last reason, it is preferable to use a model function which follows the empirical values less accurately, rather than to use a more complex function which adapts well locally to peaks on the residuals due merely to particularly noisy measurements.

After several trials, done by approximating the empirical variogram in a very sophisticated mannner, the "smoothest" and least synthetic result is that which provided us with a simple linear estimate of the empirical values of the correlogram.

In the figures below we present only those results concerning residuals in x, since the procedures are similar for both directions and the variograms in x and y are even more similar.

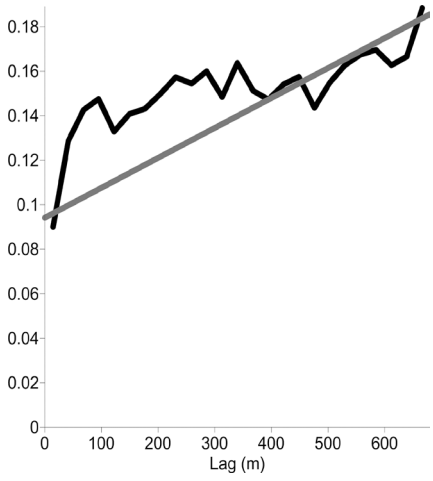


Fig. 14. Empirical variogram and residuals model in x (nugget+linear).
*Variogramma empirico e modello dei residui in x (nugget+lineare).*

Figure 14 shows empirical and model estimates of the variogram of residuals in x. We may observe that the starting "nugget" value is around 0.09 $m^2$, approximately half the global variance of these residuals. We may therefore expect a gain of about 1.4 on the values of the rms. The minimum and maximum values of these residuals in x are (–0.87, 1.13) m, and in y (–1.08, 1.15) m.

To give an idea of how these values are interpolated, we show in Fig. 15 the intentionally qualita-

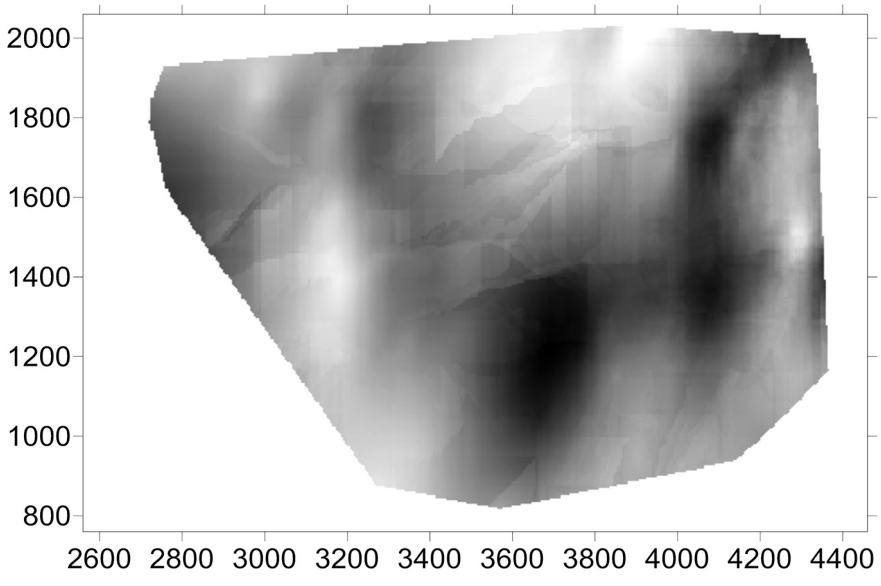


Fig. 15. Qualitative trend for the estimate of the residuals in x.
*Andamento qualitativo della stima dei residui in x.*

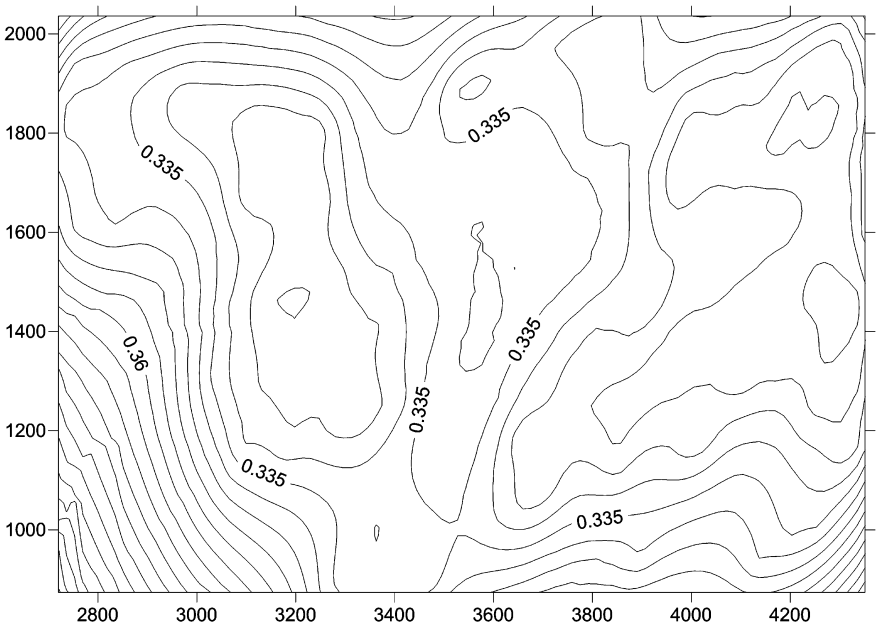


Fig. 16. Distribution of standard deviations of the estimation error in x.
*Distribuzione delle deviazioni standard dell'errore di stima in x.*

tive trend, presented in greyscale, of such estimates.

The standard deviation values of the estimate are shown in Fig. 16 and reach maximums around 0.36 m, in areas practically without measurements. As will be observed, both estimates and estimation errors follow very smooth progressions. This feature cannot be detected using other interpolators, some of which do not provide estimation errors. Even using Kriging,but with much more complex variogram functions, apparently closer to the empirical values, as in Fig. 16, the result leads to a conformation of residuals in an anything but smooth "bull's eye" (fig. 17) with estimation errors which may in places exceed 40 cm. Analogous results were obtained for deformation modelling along the y axis. (Schaffrin, 2001).

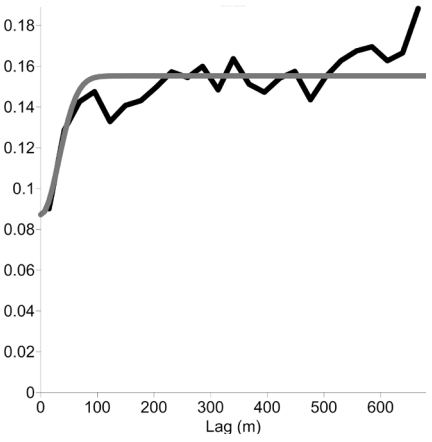


Fig. 17. Empirical variogram and residuals model in x (nugget+Gaussian).
*Variogramma empirico e modello dei residui in x (nugget+Gaussiano).*

### *3.3. Application of the two deformation components to the vector map*

It is now quite easy to understand that both the affine transformation model and the stochastic one, while derived from a subset of measurements with regard to the totality, may be applied globally to any point on the vector map. A third software package indeed, taking the digital file in CXF format and converting it into the DXF format, modifies the coordinates of each feature according to the mixed deformation model adopted. The result of this may be observed in extract form in Fig. 19.

Alternatively, the deformation model may be provided by the Cadastre in the form of a grid onto which to interpolate the overall deformation in order to correct the vector map continually.

Observing the result from a different perspective, using, that is, the DXF file format, produced by the PORI software and prior to application of the Kriging estima-

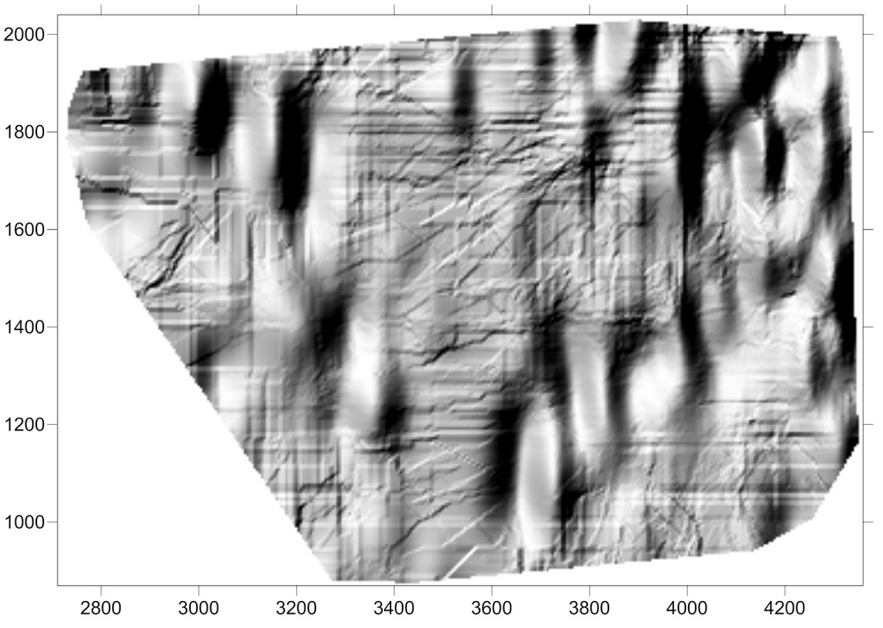


Fig. 18. "Bull's eyes" present both in the estimates and in the estimation errors which are, on average, smaller but which may locally exceed 40 cm.
*"Occhi di bue" presenti sia nelle stime che negli errori di stima che, mediamente sono inferiori ma che localmente superano i 40 cm.*


Fig. 19. Detail of verlaid cadastral vector map with the original map counter-deformed in transparency.
*Un particolare di sovrapposizione della mappa catastale vettoriale con la mappa di impianto contro-deformata in trasparenza.*

te one can make another interesting observation. Let us observe where the positions of maximum matching are located, indicated by a red circle (fig. 20) and the position of the homologous corners of the vector map after the robust affine transformation.

Enlarging certain details greatly, as in Fig. 20, it is clear that the positions of maximum matching are virtually correct: the position of the blue crosses indicating the best affine transformation is more or less correct for point 1411 but wrong for point 1412 which is only three metres to the west of point 1411 and the error is of about one metre.

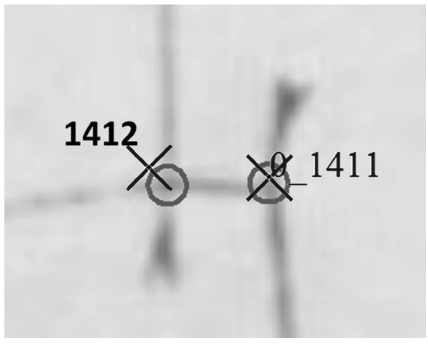


Fig. 20. A red circle marks the position of maximum matching between vector and raster while a blue cross marks that of best affine transformation of vector file.
*Con un cerchio rosso la posizione di massima correlazione tra vettoriale e raster, con una croce blu la posizione di migliore trasformazione affine del file vettoriale.*

This means that some points, even when very close, may have various deformations even up to around one metre. Such deformations are difficult to model either using a "black model", that is one with spline functions or triangular models or others, or with a stochastic approach (Crippa and Forlani, 1990).

Fortunately, not all the deformations are completely random as shown in Fig. 11 in which we may observe what is visible in Fig. 18: points very close together may have very different residuals. The stochastic approach equally gives good results, as necessary, it does not correct errors of digitization.

## 4. Conclusions

The mixed approach which involves a common deformation model for the whole map, an affine transformation and, subsequently, a stochastic approach on the remaining deformations affords numerous advantages in the recovery of the metrical information contained in the vector maps. Indeed, itt allows successive improved approximations with at least two runs and permits a first rejection of outliers during the affine transformation and a second one during the stochastic estimates.

It is also possible, where number and quality of the deformation data permit, to choose a stochastic deformation model, even more complex, which takes into account, for example, the directions of the deformation vectors. This choice was not open to us in this case due to the high value of the "random noise" of the deformations. It is the opinion of the authors that this noise is due in large part to the manual digitization of the cadastral reference maps, performed here with a precision digimeter but in a completely manual fashion, that is to say, observing each point on the map and inserting the digimeter values to each optical pointing of a corner.

The deformations would probably have been less random if the map had been acquired first with a flat scanner, but, above all, subsequently, if it had been transformed into a vector product through semi-automatic pointing allowing an automatic search for the centre or for the end of a line, or a corner of a raster feature and, if necessary, its recollimation in unfavourable cases.

## References

Bay, H., Tuytelaars, T., Van Gool, L., 2006. *SURF: Speeded Up Robust Features*. In: Leonardis, A., Bischof, H., Pinz, A. (Eds.), Computer Vision – ECCV 2006, Lecture Notes in Computer Science, 3951. Springer, Berlin, pp. 404-407.

Cina, A., De Agostino, M., Manzino, A.M., Piras, M., 2010. *Valorizzazione metrica della mappa d'impianto catastale*. Bollettino della Società Italiana di Fotogrammetria e Topografia. 1. pp. 41-54.

Cina, A., Manzino, A.M., Manzino, G., 2016. *Recovery of cadastral boun-*

*daries with GNSS equipment.* Survey Review. 48. pp. 338-346.

Cressie, N., 1993. *Statistics for Spatial Data*, Wiley, New York.

Crippa, B., Forlani, G., 1990. *Il calcolo con il metodo degli elementi finiti*. In: Ricerche di geodesia, topografia e fotogrammetria, CLUP Ed., Milano, pp. 92-113

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

Fischler, M.A., Bolles, R.C., 1981. *Random Sample Consensus: a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*. In: Communications of the ACM. 24 ( 6). pp. 381-395:

Gonzales, R., Woods, R., 2002. *Digital Image Processing*. Prentice Hall, Englewood Cliffs.

Harris, C., Stephens, M., 1988. *A combined corner and edge detector*. In: Alvey Vision Conference, 31st August-2nd September 1988, University of Manchester, pp. 147-152.

Kraus, K., 2011. Photogrammetry: *Geometry from Images and Laser scans*. Waltere de Gruyter, Berlin.

Lowe, D.G., 1999. *Object Recognition from Local Scale Invariant Feature*. In: Seventh IEEE International Conference on Computer Vision, 20-27 september 1999, pp. 1150-1157.

Oliver M.A., Webster R., 2015. *Basic steps in Geostatistics: The Variogram and Kriging*. In: Technology & Engineering Journal of the International Association of Mathematical Geology. 17. pp. 563-586

Pàroli, A., 1948. *Triangolazioni topografiche e del Catasto*. Hoepli, Milano.

Rousseeuw, P.J., Leroy A.M., 1987. *Robust Regression and Outlier Detection*. Wiley & Sons, New York.

Schaffrin, B., 2001. *Equivalent systems for various forms of Kriging, including least-squares collocation*. In: Zeitschrift für Vermessungswesen. 126. pp: 87-93.

Strang, G., 2016. *Introduction to Linear Algebra*. Cambridge Press, Wellesley, MA.

## *References from websites:*

http://www.agenziadelterritorio.it/storia-catasto-italiano.html

## *Acknowledgements*

Thanks go to Ing. Giuseppe Manzino for having provided the IT material necessary for our experimentation and for his technical advice.