AUTOMATIC GEOREFERENCING OF A HERITAGE OF OLD ANALOG AERIAL PHOTOGRAPHS

I. Cléry^a, M. Pierrot-Deseilligny^b, B. Vallet^c

^a Université Paris-Est, IGN, SRIG, MATIS, EDITE, 73 avenue de Paris, 94160 Saint Mandé, France - isabelle.clery@ign.fr ^b IGN, Laboratoire LOEMI, Université Paris-Est, 73 avenue de Paris, 94165 Saint-Mandé cedex, France - marc.pierrotdeseilligny@ensg.eu

^c Université Paris-Est, IGN, SRIG, MATIS, 73 avenue de Paris, 94160 Saint Mandé, France - bruno.vallet@ign.fr

KEY WORDS: Registration, Automation, Georeferencing, Matching, Analog, Aerial, Photography, Database

ABSTRACT:

Historical photographs become widely used in geographical and environmental applications. Their enhancement involves converting them into georeferenced data, such as orthoimages or digital models. However no ground control points are available unlike in current image processing, and many problems such as image noise, landscape modifications, perspective distortion and unknown sensor calibration prevent automatic tie-point retrieval with current orthoimages. That is why photograph georeferencing remains a manual and time-consuming task. A novel method is presented in this paper to register photographs with current topographic database using line feature matching. Indeed, geometrical considerations only let avoid high radiometric difference issues when dealing with current orthoimages. Besides topographic database use lets selecting stable through time features, such as road network and historical buildings. A multi-scale approach allows very coarse georeferencing initialization, which can be set manually by a minimum number of ground control points per image set. At each scale an iterative processing improves the line matching and the registration model estimation at the same time. Finally, building integration makes registration more reliable for off-ground objects. Results are promising as georeferencing is much improved and its estimation converges in all test cases.

1. INTRODUCTION

Historical data derived from old analog aerial photographs are more and more used in geographical and environmental applications. Historical orthoimages allow change detection of the landscape for applications such as evaluation of urban spread or highlighting of building destruction during Second World War bombarding. Historical digital surface models (DSM) allow measurement of terrain deformation or forest canopy growth. Figure 1 depicts some old aerial photographs and recent orthoimages of the same area. An important and noticeable landscape modification is mostly due to urbanization. Some information about these photographs is presented in Table 1.

At the heart of the production process is georeferencing, which implies recovering the camera position and the orientation during each photograph acquisition. Therefore, the registered historical data can directly be integrated into the geographic information systems (GIS) in order to compare to the current data. This requires a set of ground control points (GCP), which are discriminative points in the photographs and known as terrain coordinates. Ground control point database are usually available in the present aerial photographs process. However, due to the unavailability of the GCP database, the operators have to visually search for the homologous objects in the old scanned photographs and present orthoimages.

They also use current digital terrain model (DTM) to get altitude coordinates, as no historical DSM or DTM are available. This method is time-consuming as well as unreliable. Indeed a first concern is that DTM use involves getting only onground GCP. Then registration parameter (focal length in particular) are coarsely estimated and off-ground objects such as buildings are badly registered. This is not desirable for both historical DSM computation of old registered overlapping aerial photographs set and complete orthorectification of old photographs which include building rectification (using historical DSM information).

This is problematic for historical DSM computation from a set of old registered overlapping aerial photographs, or for

complete orthorectification of old photographs including building rectification using this historical DSM information. The second concern is the difficulty to select reliable GCP in the images of different dates, due to: important landscape changes during the very long time gap (30 to 90 years), seasonal and every day effects (vegetation modification, shadows and moving cars), which are visible at such resolution, i.e. 20 to 30 cm, perspective effects on off-ground objects, sensor and scanner noise, and film defects, such as scratches, holes and tears.

This issue not only affects the operator work but might also cause the failure of the algorithms that automatically extract homologous points from images.

In addition to this, control features must be selected on stable through time objects. Unfortunately everything is likely to be changed. Annual rise of rivers and lakes, building construction, modification or destruction, vegetation growth, crop renewal, road enlargement or deviation, cross-road enhancement with roundabout, fields merging, urban spreading over land and forests, rail track construction or neglecting are some examples. However some permanent structures remain same, such as historical buildings, road network and rock appearance in natural areas.

Therefore our goal is to find an automatic method to register old scanned photographs onto current data with weak initial georeferencing assumptions, such as a place name is usually indicated on photograph border or box. Off-ground object registration should be taken into account. Target dataset is limited to aerial subvertical images of humanly-modified landscapes that contain roads and/or buildings. This excludes entirely natural landscapes and slanted shootings but applies to the majority of heritage photographs.

The following part describes related work on the subject. The proposed method is presented in Section 3. In Section 4, it is tested on different data set to evaluate its reliability and result precision. Finally interest and limitations of the method are discussed and perspectives are put forward.



Figure 1. Old aerial photograph (left) and current orthoimage (right) of the same areas.

From top to bottom : Agde, Nantes_1, Nantes_2, Saint-Mandé. Main modifications are the important urbanization and roundabouts for Nantes_1, the shipyard replacement by buildings and the river covering (bottom of image) for Nantes_2, the motorway construction (top left of image) and the field urbanization (right of image) for Saint-Mandé, and the modification of most buildings in all cases.

2. RELATED WORK

Historical photographs are objective witnesses of our surrounding at a precise time that may go back until a century ago unlike satellite images. That is why they become more and more important in many applications. Thus (Wiedemann, 2000) reconstructs 3D models of destroyed buildings of Berlin. Terrain deformation can also be measured on georeferenced aerial images. (Walstra, 2004) computes vertical and horizontal displacement maps from digital terrain models at different dates to study correlation between landslides evolution and climatic conditions. (Ayoub, 2009) makes a terrain displacement map after to an earthquake. One of the key step, i.e. photograph enhancement, is handled by (Redweik, 2010) in which they register old aerial photographs of Portugal and displays the result on a graphical application.

Photograph georeferencing represents the main technical bottleneck for process automation due to the unavailibility of the ground control points. The first idea to solve this problem is to adapt usual automatic methods that extract tie-points between photographs and current (or more recent) photographs or orthoimages. For example, (Liu, 2006) refines multi-date aerial image registration (one year gap only). In this case, homologous points are extracted by cross-correlation and filtered by local coherence of translation error. These methods include point of interest detection algorithms, such as Scale-invariant feature transform (SIFT), Speeded-up robust features (SURF) or Maximally Stable Extremal Regions (MSER) (see Tuytelaars, 2008 and Mikolajczyk, 2005a for a review). Point of interest are associated with a local descriptor such as the SIFT descriptor for point matching (see Mikolajczyk, 2005b for a review). These methods include also area-based matching methods such as cross-correlation or mutual information (Maes, 1997). However, they fail to extract reliable tie-points due to both important difference of time and change of sensor between data. Few methods are proposed to deal with the too little number of tie-points. (Mizotin, 2010) deals with low overlapping aerial images by separating the estimation of the different pose parameters (firstly yaw, secondly 2D translation, then all parameters together). (Müller, 2012) uses multi-scale approach to detect more tie-points in multi-date satellite images, by local affinity constraint computed from previous step tie-points and cross-correlation. Even so, these methods do not deal with both problems and are often based on the existence of the reliable initial tie-points or a quite precise registration, which is not the case of our data.

Similar algorithms deal with other kinds of features, such as contours and lines. The best-known detector is (Canny, 1986)'s applies a connectivity aware hysteresis threshold to the local maxima of the gradient norm in the gradient direction. A line estimation step may be added to get the line features. (Martin, 2004) segments the images using intensity gradient, colour gradient (in case of a RGB image) and texture gradient. Line segment detector (LSD) algorithm (Grompone, 2012) uses an *a* contrario approach to merge areas of pixels with the same gradient direction. (Desolneux, 2000) meaningful alignments algorithm exhaustively scans all image line directions and selects segments which contain the most significant number of pixels that have the same gradient direction as the segment (the meaningfulness depends on the segment length and on the number of pixels of correctly oriented gradient). Moreover this algorithm is robust to alignment interruptions.

Extracted line features are then matched with other images or with Light detection and ranging (LIDAR) or Synthetic aperture RADAR (SAR) data. (Borgefors, 1988) uses a multi-scale approach to register an image on an other using linear features. At each scale, registration is refined by correction trials of every registration model parameter. Then a voting approach, based on mean Euclidean distance of extracted pixels to projected lines, is used to select the new parameter values. Besides the initial registrations are set from all possible values of parameters and progressively filtered at each scale. However, mean Euclidean distance strongly depends on line density in image, that is why this method fails in case of important over- detections and under-detections. In the context of different date image matching, (Vassilaki, 2012) selects the lines that are matched by closest centroid and line ends. Then line break points are automatically extracted and used as GCP. An iterative approach enables to match points and estimate projection parameters. (Li, 1995) registers multi-sensor (optical or even SAR) data. For that purpose, contours are detected and coded by their internal successive orientations, then contours are matched by principal axis length coherence, and contour points by maximization of cross-correlation of contour coding strings. (Oh, 2012) registers satellite images on LIDAR data. In this context, contours are extracted with Canny detector, then a grid of GCP is computed by a cross-correlation-like measure on binary data (contour or not contour pixel) and filtered by distance between measure maximum and following values. However, contour-based approaches suppose that separated full contour lines can be extracted and that break points or contour points are stable features. This is rarely the case, especially in urban areas or with images of different scale and noise level.

Line features are also fully compatible with image registration on topographic database or map. (Barsi, 2004) retrieves a line network in an image from a graph model, a point set extracted by image threshold and a fine initial position. Each graph node is iteratively moved to the closest extracted point and each nonmatched node is moved accordingly to the connected node displacement. (Roux, 1993) automatically registers coarsely georeferenced satellite images on a map using urban conglomeration and cross-roads as GCP. In this case, detection is made by morphological filtering and linear features detection. Then registration incorporates Random sample consensus (RANSAC, see Fischler, 1981), which is a voting approach, followed by least square minimization. (David, 2003) registers iteratively an image onto a topographic line database. For that purpose, topographic line ends are projected on the image and matched with the nearest linear feature detected in image with Canny detector. Then an orthography is estimated by weighted least square method in order to minimize point-to-line distances. Weight values are the sum of squared point-to-line distances and line orientation difference. (Frueh, 2004) registers slanted aerial images in a similar fashion, except that initial matches are exhaustive and not only closest neighbour, and that weights include line lengths. Initial parameters are obtained either with manual GCP, or by an exhaustive research over all parameters possible value interval. (Chen, 2006) registers a topographic database on a high-resolution orthoimage using cross-roads as GCP. In this case, roadsides are extracted with Canny detector and road area by image threshold. Then crossroad centre position is precisely set by binary cross-correlation with a template made from the database node form (connected road orientations and widths). The major shortcoming of all these methods is the need of a quite fine initial georeferencing. Another issue is the extraction of already well-filtered features of interest for data matching, with radiometric criteria for (Barsi, 2004), (Chen, 2006) and (Roux, 1993) that depend much on photograph features.

Finally (Stoica, 2004) extract road network with Gibbs point process while (Ruskoné, 1996) uses road following from seeds with decision tree (selection of the path of lowest radiometric variance) and object supervised classification to identify and pass over obstacles. However such methods are based on the strong radiometric hypothesis, such as road colour uniformity and high road limit gradient. That is why they usually fail when there are obstacles, such as shadows or cars. Supervised classification is unusable in our case because manual GCP selection would be as fast.

3. AUTOMATIC METHOD FOR OLD PHOTOGRAPH GEOREFERENCING

3.1 Data requirements

As depicted in Figure 1, the first observation made while comparing old aerial photographs to recent images of the same area (photographs or orthoimages) is that the more stable through time and the best visually recognizable features are linear (in particular road networks and buildings). Indeed linear features provide more geometric information than point feature. On the contrary, point features can barely be discriminated because local invariant descriptors of the literature do not manage many modifications of point neighbourhood. These modifications may be perspective distortion of roof corners or non-uniform radiometric variation of objects (for example crop change in fields). That is why we focus on linear feature detection and matching, and try to get rid of radiometric descriptor analysis. We consider available current data which include aerial photographs, orthoimages, digital terrain models, maps and topographic database (TopoDB). Starting from that, it can be noted that linear feature detection onto current images (photographs or orthoimages) creates inevitably over- and under-detection issues in addition to those from old photographs. An example is shown in Figure 2. In this case, LSD algorithm (Grompone, 2012) detects field furrows in the current orthoimage and many less in the old photograph due to image noise. Moreover, some main roads which are detected in the old photograph do not match any line of the current orthoimage due to new tree rows all along these roads.

On the contrary, linear features extracted from TopoDB are exhaustive and precisely positioned (they are extracted from georeferenced aerial photographs by stereoscopy). TopoDB is provided with useful properties, such as point altitude, which avoids extraction by DTM interpolation and the resulting approximation. Some examples are shown in Figure 4 and Figure 6.



Figure 2. Illustration of over-detection problem when using an orthoimage as reference

Left : old photograph; right : current orthoimage. Lines are detected with LSD algorithm.

Hence our method requires the existence of such a TopoDB that contains roads (stored as a 3D polyline for central axis position and road width) and buildings (stored as a 3D point polyline that correspond to building envelop). It also requires a coarse initial georeferencing for every image in order to extract all TopoDB features in the photograph footprint. Several methods can produce such an initial georeferencing for a lower cost. As aerial photographs are usually a part of an aerial mission, this initial georeferencing may be an index table. This can be achieved by an operator for the whole photograph set who positions on a map only each band end photographs. Other photograph positions are then interpolated. This is the case for our data. The second option is to compute tie-points between photographs of the set, followed by estimating a global relative orientation of camera poses, and finally to get an absolute global orientation using only three ground control points. Besides this option enables the computation of the camera calibration while estimating relative orientations. Therefore it takes an eventual distortion into account in the following processing. The third option is to use a previously georeferenced photograph of the same area, which was shot at a date close to the query photograph date. Hence homologous points can be selected automatically. In this case, the query photograph should be georeferenced using current database in order to avoid georeferencing error propagation.

The proposed approach consists in registration of the each photograph independently by estimating a full projection function without distortion. This enables the georeferencing of single photographs, but it can also be a first process before computing a global bundle adjustement in the case of a photograph set. Indeed ground control features can straightfully be extracted from the results.

Consider a ground point of the TopoDB located at coordinates X, Y and Z, which is projected on the image coordinates x and

y. Note also a_i , b_i , c_i , d_i , *i*=1,2,3 the eleven parameters of the projection function to estimate. Then the full projection is defined in Equation 1.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a_1 X + b_1 Y + c_1 Z + d_1}{a_3 X + b_3 Y + c_3 Z + 1} \\ \frac{a_2 X + b_2 Y + c_2 Z + d_2}{a_3 X + b_3 Y + c_3 Z + 1} \end{bmatrix}$$
(1)

3.2 Projection estimation from line features

At this step, a set of segments $\{l_i\}_{i=1}^N$ is assumed to be extracted from the photographs to be processed. More detail about the line feature detection is presented in Part 3.3. Every line l_i is defined by its direction unit vector V_i and its signed distance to the origin y_i .

In addition, a set of segments $\{L_j = [P_{1,j} P_{2,j}]\}_{j=1}^M$, where $P_{1,j}$ and $P_{2,j}$ are line L_j ends, is assumed to be selected from topographic database by considering the approximate footprint of the photograph.

Our method to estimate the projection function of Equation 1 is based on (Frueh, 2004). More precisely, L_j is projected on the image by an initial solution. Then every projected line is matched with image lines. The three criteria presented in Figure 3 are computed: relative overlap (further referred as r), mean distance in pixels (referred as d) and orientation difference in rad (referred as or) between the two segments. The match is selected if relative overlap is above a threshold s_r , mean distance is above a threshold s_{or} . The three threshold initial values depend on the expected precision of the initial solution. Likewise multiple matches are allowed. Matches form a set M

Then the projection function of Equation 1 is estimated by an iterative weighted least square estimation. If *w* refers to a weighting function, Π to the projection function and \times to the 2D cross-product, then the energy to minimize is given in Equation 2.

$$\sum_{(i,j)\in M} w(i,j) ([V_i \times \Pi(P_{1,j}) - y_i]^2 + [V_i \times \Pi(P_{2,j}) - y_i]^2)$$
(2)

The strength of this equation is its independence from line end positions as line ends are not supposed to match. Besides it enables the computation of all model parameters at the same time.

Match weights $\{w(i, j)\}_{(i, j) \in M}$ for Equation 2 are set from the criteria and thresholds previously depicted in Figure 3 and are defined in Equation 3.

$$w(i,j) = \frac{(r-s_r)}{s_r} \frac{(\cos(\text{or}) - \cos(s_{\text{or}}))}{\cos(s_{\text{or}})} \frac{(s_d - d)}{s_d} \|\Pi(P_{1,j})\Pi(P_{2,j})\|$$
(3)

where $\|.\|$ is the euclidean distance in image frame.

A little improvement compared to (Frueh, 2004) is the introduction of the overlapping threshold for matching selection and for Equation 2 weight. Indeed this avoids matching each road segment with the whole succession of the detected segments and reciprocally. Note that the possibility of the negative value of threshold in the first iterations corresponds to a longitudinal distance between the matched segments.

At each iteration, matching is constrained by lowering the three criteria thresholds to enhance registration precision, and initial solution is updated from the previous iteration solution. This method is further referred to as line-based SoftPOSIT algorithm (David, 2003).



Figure 3. Line matching criteria

3.3 Coarse registration of subsampled images

In the case of old aerial photographs, the coarse initial georeferencing makes line-based SoftPOSIT algorithm (David, 2003) impractical. Indeed, the most weighted matches are false, so that the estimated solution is always false. This is more visible in urban areas where road networks look have an array structure; hence the streets are easily matched with parallel ones. (David, 2003) and (Frueh, 2004) fix this issue by using a set of initial georeferencings so that at least one should fall in the correct solution attraction interval. In practice, this method is unusable for urban area because there is a local minimum for every street interval and building alignment interval. The computation time would be totally incompatible with a production context. Moreover, whenever the line detection algorithm used, a high number of over-detections is obtained due to the presence of building alignments, building shadows or field furrows that look like road sides. An example in a urban area is presented in Figure 4. A few field furrow over-detections are visible on the old photograph of Figure 2. That is why a voting method such as Random sample consensus (Fischler, 1981) is preferred. In this way, topographic database lines are projected on the image and matched with neighbouring detected lines. Then an iterative random selection of a minimum number of matches enables computation of projection parameters. The solution is finally selected by counting the number of the matched inliers explained by estimated projection.



Figure 4. Illustration of pavement detection difficulty Left : detected lines in image with LSD algorithm, most of them lie on building shadows or along buildings

Right : projection of TopoDB on the image (road axis are in blue, pavements are in red,and buildings are in green). Pavements are a little lighter than streets and barely visible

Another issue is risen by the unreliable road side detection position. Terrain pavement position is deduced from central axis position and road width, which has only metric accuracy and may have changed through time. The detected line in the image is based on the gradient maxima extraction, which often appears along building limits, building shadow limits, tree rows or lane markings (pavements are barely visually detectable, especially when shadows or trees hide them, which is illustrated in Figure 4). Detected lines are close to the road side but not often exactly on it. That implies a biased registration result, particularly with a voting approach that selects badly spread matches over the image. That is why the initial coarse registration is made on the subsampled image in order to reduce the number of matches. Moreover TopoDB is filtered by road width criteria because only main roads are detected. Subsampling scale is chosen such that the most of the road widths are less than a pixel in image. At this step, lines are detected in image using a meaningful alignments detection algorithm (Desolneux, 2000) which is robust to noise and alignment interruption. In addition, subsampling decreases computational complexity of exhaustive alignment research. This step is illustrated in Figure 5. The whole algorithm is summarized in Algorithm 1.

A coarse to fine approach is used to estimate the model. At this step a simple 2D affinity is estimated to refine initial georeferencing. This simplification is justified by the use of the subvertical aerial photographs. Using topographic point coordinates as defined in Equation 1, the six parameters $a_{i,j}$, b_i , i, j=1,2, of affinity can be estimated with Equation 4.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$
(4)

The limited number of parameters helps to restraint the number of trials, and the probability that selection contains a false match. The number of trials is empirically set to 1000, which corresponds to an inliers rate of 1/6 and a probability of success of 0,99.

Obtained coarse registration is then refined by previously presented line-based SoftPOSIT algorithm (David, 2003). However relief distortion is often not visible at such scale. Hence degenerated configurations of the road network are checked by computing the mean network plane and checking mean distance between the road points and the closest plane points after projection by initial georeferencing. Indeed, this is an apporximation of the error in image between point projection with a homography and projection with Equation 1. A distortion under 2 pixels indicates that the terrain is too planar, i.e. terrain distorsion in image is not visible. In this case, an homography is estimated instead. The eight parameters a_i , b_i , c_i , i=1,2,3 of homography are estimated with Equation 5.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{a_1 X + b_1 Y + c_1}{a_3 X + b_3 Y + 1} \\ \frac{a_2 X + b_2 Y + c_2}{a_3 X + b_3 Y + 1} \end{bmatrix}$$
(5)

3.4 Full resolution registration with roads

As previous step provides a good approximate registration, linebased SoftPOSIT algorithm (David, 2003) can then be used to estimate the full projection defined in Equation 1. This step is shown in Figure 6. Road areas are clearly visible at full resolution, hence roads width w_i are taken into account by a modification of the distance criterion (see Figure 3) which allows multiple matches (for both roadsides). The criteria threshold is also adapted to possible road width and road detection uncertainties with the additional term $w_i/4$. Indeed the widest road sides have less stable widths, are more often modified through time, more often masked by tree or car rows, and are lined with larger pavements that are not always detected (refer to Figure 4 for road side detection approximation). These updates are presented in Equation 6.

$$\left| d - \frac{w}{2} \right| \leq s_d + \frac{w}{4} \qquad (6)$$

It can be noted that roads are not split into two distinct polylines describing roadsides because starting position of road projection is not always inside road area, so closest projected roadside bonds to the wrong image line, which induces a shift equal to road width. Besides degenerated configurations of road network are checked.

Line segment detector (LSD, Grompone, 2012) algorithm is chosen to detect lines in image because of its speed compared to meaningful alignments, its precision as it does not fill holes and because it directly provides line features unlike Canny-like algorithms (Canny, 1986). Indeed road network and building limits are mostly straight, so it does not miss much information in case of winding roads.

In the end, topographic database roads are projected on the photograph inside their image area. However, road axis is not centred in road area due to line detection method, which does not extract exactly road pavements. Furthermore, off-ground object projections are shifted as depicted in Figure 6 bottom left, so next step consists in introducing off-ground features to refine the result.



Left : initial projection of TopoDB road network on the image Right : meaningful alignments detected in image



Left : TopoDB projection on the image with an affinity estimated by RANSAC algorithm (main roads are in red) Right : TopoDB projection with an homography estimated by line-based SoftPOSIT algorithm

Figure 5. Registration of subsampled image from Agde dataset (presented in Figure 1 and Table 1)

3.5 Precise off-ground registration with buildings

This step is similar to registration with road network, except that detected lines correspond exactly to building limits, so no biases are expected on registration estimation. The comparison with previous step can be seen in Figure 6 bottom right. However, the huge number of buildings and so of line matches makes implementation more complex. That is why a filtering is introduced by considering local coherence of building line matches. It relies on two observations. Firstly correct projection of every building is quite close to the position given by previous step result (discrepancy depends on road width in image). Secondly error corresponds locally to a 2D translation at constant altitude. For every building, two non-parallel line matches allow computation of translation error parameters. Hence the number of inliers among neighbouring (both planimetrically and altimetrically) building line matches qualifies robustness of local correction. Thus only the inliers of best translation are kept in each area, which decreases algorithm complexity.



Left : Initial projection of TopoDB road network on the image (with solution obtained at previous step)



Left : TopoDB projection with model defined in Equation 1 estimated from roads only. Building projections on the top of the image do not well match roof shapes. Right : TopoDB projection with model defined in Equation 1 estimated from both roads and buildings

Figure 6. Registration of full-resolution image (crop) from Agde dataset (presented in Figure 1 and Table 1).

ALGORITHM 1

Data : *I* input grayscale aerial photograph, Π^0 coarse initial projection, L^R a set of road segments, L^B a set of building limit segments, both from a topographic database **Result** : Π final projection

begin

 I^{dz} = Subsampling (I) l^{dz} = MeaningulAlignments (I^{dz}) \rightarrow (Desolneux, 2000) M^{dz} = Compute Matches ($\Pi^0(L^R)$, l^{dz}) → Sec. 3.2 Π^{dz} =RANSAC (M^{dz}) → Eq. 4, 2 b^{dz} = CheckRoadNetworkRelief (L^{R} , Π^{0}) → Sec. 3.4 $\Pi^{dz} = \text{SoftPOSIT} (\Pi^{dz}, L^{R}, l^{dz}, b^{dz}) \rightarrow Eq. 5 \text{ or } 1, 2, 3$ l = LSD(I) \rightarrow (Grompone, 2012) $b = \text{CheckRoadNetworkRelief}(L^{R}, \Pi^{0})$ → Sec. 3.4 $\Pi = \text{SoftPOSIT} (\Pi^{dz}, L^{R}, l, b)$ \rightarrow Eq. 1 or 5, 2, 3, 6 S^{B} = Segmentation ($\Pi(L^{B})$) → Sec. 3.5 $M = \text{ComputeMatches} (\Pi(L^B), l)$ → Sec. 3.2 foreach i do → Sec. 3.5 $S_i^{B'}$ = Filtering (S_i^{B} , I, M) $L_i^{\scriptscriptstyle B}$ ' = arg ($L_i^{\scriptscriptstyle B}$ | $\Pi(L_i^{\scriptscriptstyle B})$ = $S_i^{\scriptscriptstyle B}$ ') end $L^{^{R}} <= L^{^{B}}'$ $\Pi = \text{SoftPOSIT}(\Pi, L^R, l)$ \rightarrow Eq. 1, 2, 3 end

Finally, equations on roads are kept and adapted to split topographic database lines into two side lines, as detected segment position from central axis is now known. However road equation weights (defined in Equation 3) are decreased to take into account the uncertainty on pavement detection.

4. EXPERIMENTAL EVALUATION

4.1 Data

The French National Photograph Library is progressively making its heritage available to the general public by digitizing and making it available on the Internet. Thus images from different kind of landscapes of the French territory (urban areas, rural areas and urbanized former rural areas) have been selected and are shown in Figure 1.

Initial georeferencing is given by an index map (IM) for each photograph set. Its precision is said to be equal to one hundredth of analog scale (in meters). This corresponds to operator pointing precision for each photograph positioning. Parameters for the different steps are listed in Table 4.

The full method process is applied to one or several images of every set. Data specifications are detailed in Table 1.

Location	Agde	Nantes_1	Nantes_2	Saint- Mandé
Year	1968	1923	1923	1936
Scale	25000	10000	10000	11500
Landscape	Rural with towns	Former rural, urbanized	Urban (city centre)	Urban (city suburbs)
Scan resolution	1200 dpi	750 dpi	750 dpi	750 dpi
IM precision	250 m 395 px	100 m 274 px	100 m 289 px	115 m 275 px

Table 1. Data specifications of the old photographs presented in Figure 1

4.2 Results

A set of check points is measured for each photograph. They are homologous points between topographic database and image, and are selected at cross-roads for one-half and at building corners for the other half. The quality of georeferencing is given in Table 2 and Table 3. Table 2 shows image residuals, which are distances between image coordinates and coordinates of ground points projected on the image. Table 3 shows terrain residuals, which are distances between ground coordinates and vector of back projected image points.

It has to be noted that the check points are themselves soiled by a pointing error which is more important than errors met with current images. It is due to the important noise in images, to the difficulty to retrieve the same objects in different time data between a completely changed surrounding, and to the fact that invisible modifications may happen, such as road enlargement or translation, or building raising. The projection function of Equation 1 is estimated from check-points, to check both projection model relevance and check-points consistency. Doubtful check-points are then removed or replaced when possible. Check-points precision is set to about 3 pixels. Besides topographic database precision, reliability and completeness are not checked or discussed here, but its precision is said to be about one meter. In addition, when manually processing historical photographs, GCP are usually removed when their residuals are higher than one meter to match final orthoimage quality requirement. Hence this value is taken as goal precision.

Although this method involves several iterative computations at different scales, it takes about 2 minutes per image to complete

all the steps, including line detection, topographic data extraction and multi-scale registration. Moreover it can run independently on several images at the same time depending on the number of available processors or computers. Considering that it is fully automatic except for coarse initial georeferencing setting, the total time-saving for an operator is very important. Indeed current historical orthoimage production requires selection of one ground control point per photograph to ensure enough correct GCP.

Step	Check- points	Agde	Nantes_1	Nantes_2	Saint- Mandé
IM	all	737.30	56.48	135.14	93.53
	cross-roads	739.88	53.99	133.17	96.29
	buildings	735.01	58.98	136.91	90.47
Step 1	all	53.67	63.07	32.92	45.40
	cross-roads	55.54	71.02	32.83	45.99
	buildings	37.24	55.13	33,00	44.75
Step 2	all	4.37	26.02	18.70	11.84
	cross-roads	4.93	27.19	17.52	11.00
	buildings	3.87	24.84	19.78	12.77
Step 3	all	2.75	22.75	6.07	4.51
	cross-roads	3.55	20.13	6.58	4.89
	buildings	2.04	25.36	5.62	4.09

Table 2. Image residuals in pixels at each step for the different kinds of check-points

IM indicates index map measured precision; step 1 is subsampled image registration by an homography (see Figure 5); step 2 is full-resolution image registration using road network only (see Figure 6 bottom left); step 3 is final registration using both buildings and roads (see Figure 6 bottom right)

Step	Check- points	Agde	Nantes_1	Nantes_2	Saint- Mandé
IM	all	432.36	20.66	43.75	37.22
	cross-roads	429.93	19.75	42.70	38.13
	buildings	434.53	21.58	44.70	36.20
Step 1	all	34.08	24.28	11.94	18.30
	cross-roads	35.25	27.06	11.89	18.60
	buildings	33.31	21.50	11.98	17.96
Step 2	all	2.56	10.11	6.84	4.24
	cross-roads	2.88	10.37	6.38	3.94
	buildings	2.27	9.86	7.25	4.57
Step 3	all	1.57	3.25	1.52	1.77
	cross-roads	2.02	5.49	1.75	1.93
	buildings	1.16	1.01	1.31	1.60

Table 3. Terrain residuals in meters at each step for the different kind of check-points (steps are the same as in Table 2)

Step	S _d	s_d inliers	S _{or}	s ₀r inliers	s _r	s _r inliers
RANSAC	>E/100px	E/1000px	30°	10°	- <i>s</i> _d px	0,5 %
Step 1	E/500 px	dz px	10°	5°	0,5 %	0,75%
Step 2	2*dz px	dz/5 px	5°	1°	0,5 %	0,9 %
Step 3	dz/5 px	1 px	5°	1°	0,5 %	0,9 %

Table 4. Parameters at each step

E is the analog scale, dz the subsampling scale. Distance due to rotation error is added to RANSAC s_d .

4.3 Discussion

The first constatation is that each step of the method significantly improves registration precision. Proceeding step by step, it can be noted that index map precision is not at all what was expected. Indeed no check-points of Agde set get a residual under 395 px, whereas all check-point residuals of Saint-Mandé are under 275 px. This may be explained by the fact that Saint-Mandé photograph is at a band end whereas Agde photograph is the fourth of the band which means that its position in the IM was computed by interpolation.

Subsampled image registration precision should be divided by sampling scale (about 1/20) to correspond to detected line precision.

Full-resolution image registration with road network only obtains a precision that is compatible with mean road width in the image, except for Nantes_1. It should be noted that an homography (Equation 5) was estimated for Nantes set, unlike Agde and Saint-Mandé sets. This explains why residuals decrease much between step 1 and step 2 for Agde and Saint-Mandé.

Finally, registration with both buildings and roads enhance all kind of check-point residuals (actually a little more building check-points), which corresponds to left uncertainty at previous step when little reliable line pavements were used.

The second constatation is that the algorithm converges in all scenarios. This is a quite surprising result for Nantes_1 where check-points were particularly hard to find. Indeed most cross-roads have be turned into roundabouts (see Figure 1), and only a few buildings existed before complete urbanization of the area. Moreover, most of them were destroyed and rebuilt. In this particular case, line feature extraction and matching approach is very interesting because roads were nearly not modified, unlike cross-roads, which makes this method highly promising.

The third constatation is that all residuals are still high, and an additional filtering needs to be made on building line. First improvement may be to correct image distortion before processing. However, this does not justify the whole error value. If the entire set of photographs is available, a second improvement is to integrate last step equations in a complete bundle adjustment, so that multiple images information will help filter line mismatches. This allow photograph tie-point addition in data, and calibration parameters addition in unknown set. Photographs without line networks can also be registered in this way. A third improvement is to enhance filtering of building line matches. Indeed, there are still a lot of outliers after final registration, which is due to the difficulty for the algorithm to distinguish modified buildings (actually it is quite difficult even for an operator). 3D information brought by bundle adjustment and radiometric information are possible solutions to explore.

5. CONCLUSIONS

A method to automatically georeference old analog photographs was presented. Digitized photographs are registered one by one with a topographic database as reference, which contains road network and building limits in the form of 3D polylines. A multi-scale approach is used to deal with very coarse initial georeferencing. At each scale, iterative processing matches detected lines in images with topographic lines and estimates registration parameters, based on three criteria: distance, orientation difference and overlap between matched lines. Data selection, line extraction, registration model and parameter computation are adapted to current scale. The high number of buildings is taken into account by a local geometric filtering.

The algorithm is tested on images, which represent the different kind of landscapes (rural, urban and suburban areas) that are subject to very important changes. Results are highly promising as georeferencing precision measured on check-points is much improved, and because the registration converges in all scenarios even when ground control point retrieval is an extremely difficult task for operators. Moreover topographic building limits are useful data in georeferencing as they lower check-point residuals and make registration more reliable on off-ground objects. This method can also be used for recent photograph georeferencing when automatic algorithms for GCP extraction fail.

However, results do not reach expected quality requirements yet. Further improvements include registration precision enhancement by filtering building line matches and by using statistical discrepancy lowering provided by bundle adjustment on whole data set when available.

6. ACKNOWLEDGEMENT

Authors acknowledge Prof. Julie Delon of Université Paris Descartes for important contribution to the work, especially in line detection bibliography and for detailed paper review, Neelanjan Bhomwik of Laboratoire MATIS, IGN, for full paper review, and anonymous reviewers for their useful corrections.

7. REFERENCES

Ayoub, F., Leprince, S., et Avouac, J.-P.. Co-registration and correlation of aerial photographs for ground deformation measurements. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2009, vol. 64, no 6, p. 551-560.

Barsi, A. Object Detection using Neural Self-Organization. In : *Proceedings of the XXth ISPRS Congress*, Istanbul, Turkey. 2004, XXXV (B3), 366-371

Canny, J. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1986, no 6, p. 679-698.

Chen, C.-C., Knoblock, C. A., et Shahabi, C. Automatically conflating road vector data with orthoimagery. *GeoInformatica*, 2006, vol. 10, no 4, p. 495-530.

David, P., Dementhon, D., Duraiswami, R., et al. Simultaneous pose and correspondence determination using line features. In : *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on.* IEEE, 2003. p. II-424-II-431 vol. 2.

Desolneux, A., Moisan, L., et Morel, J.-M. Meaningful alignments. *International Journal of Computer Vision*, 2000, vol. 40, no 1, p. 7-23.

Fischler, M. A. et Bolles, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 1981, vol. 24, no 6, p. 381-395.

Frueh, C., Sammon, R., et Zakhor, A. Automated texture mapping of 3D city models with oblique aerial imagery. In : *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on.* IEEE, 2004. p. 396-403.

Grompone Von Gioi, R., Jakubowicz, J. Morel, J.-M., et al. LSD: a line segment detector. *Image Processing On Line*, 2 (2012), pp. 35–55. http://dx.doi.org/10.5201/ipol.2012.gjmr-lsd

Li, H., Manjunath, B. S., et Mitra, S. K. A contour-based approach to multisensor image registration. *Image Processing, IEEE Transactions on*, 1995, vol. 4, no 3, p. 320-334.

Liu, D., Gong, P., Kelly, M., et al. Automatic registration of airborne images with complex local distortion. *Photogrammetric engineering and remote sensing*, 2006, vol. 72, no 9, p. 1049

Maes, F., Collignon, A., Vandermeulen, D., et al. Multimodality image registration by maximization of mutual information. *Medical Imaging, IEEE Transactions on*, 1997, vol. 16, no 2, p. 187-198.

Martin, D. R., Fowlkes, C. C., et Malik, J. Learning to detect natural image boundaries using local brightness, color, and texture cues. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2004, vol. 26, no 5, p. 530-549.

Mikolajczyk, K. et Schmid, C. A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2005, vol. 27, no 10, p. 1615-1630.

Mikolajczyk, K., Tuytelaars, T., Schmid, C., et al. A comparison of affine region detectors. *International journal of computer vision*, 2005, vol. 65, no 1-2, p. 43-72.

Mizotin, M., Krivovyaz, G., Velizhev, A., et al. Robust matching of aerial images with low overlap. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2010, vol. 38, p. 13-18.

Müller, R., Krauß, T., Schneider, M., et al. Automated georeferencing of optical satellite data with integrated sensor model improvement. *Photogrammetric engineering and remote sensing*, 2012, vol. 78, no 1, p. 61-74.

Oh, J., Lee, C., Eo, Y., et al. Automated georegistration of highresolution satellite imagery using a RPC model with airborne LiDAR information. *Photogrammetric Engineering & Remote Sensing*, 2012, vol. 78, no 10, p. 1045-1056.

Redweik, P., Roque, D., Marques, A., et al. Triangulating the Past-Recovering Portugal's Aerial Images Repository. *Photogrammetric engineering and remote sensing*, 2010, vol. 76, no 9, p. 1007-1018.

Roux, M., Lopez-Krahe, J., et Maître, H. Recalage image SPOT/carte routière. *International Archives of Photogrammetry and Remote Sensing*, 1993, vol. 29, p. 384-384.

Ruskoné, R. Extraction automatique du réseau routier par interprétation locale du contexte: application à la production de données cartographiques. 1996. *PhD thesis*. Marne-la-Vallée.

Stoica, R., Descombes, X., et Zerubia, J. A Gibbs point process for road extraction from remotely sensed images. *International Journal of Computer Vision*, 2004, vol. 57, no 2, p. 121-136.

Tuytelaars, T. et Mikolajczyk, K. Local invariant feature detectors: a survey. *Foundations and Trends*® *in Computer Graphics and Vision*, 2008, vol. 3, no 3, p. 177-280.

Vassilaki, D., Ioannidis, C., Stamos, A. Recovery of the Geometry of Historical Aerial Photos Associating Self-Calibration with Ground Control Linear Features. In : *EARSeL Workshop Proceedings: First International Workshop on Temporal Analysis of Satellite Images*. 2012. p. 202-207.

Walstra, J., Chandler, J. H., Dixon, N., et al. Time for changequantifying landslide evolution using historical aerial photographs and modern photogrammetric methods. In : Proceedings 20th Isprs Congress. 2004. p. 475-480.

Wiedemann, A., Hemmleb, M., Albertz, J. Reconstruction of historical buildings based on images from the Meydenbauer archives. *International Archives of Photogrammetry and Remote Sensing*, 2000, vol. 33, no B5/2; PART 5, p. 887-893.