AUTOMATIC EXTERIOR ORIENTATION OF AERIAL IMAGES IN URBAN ENVIRONMENTS

C. Drewniok, K. Rohr
Fachbereich Informatik, Universität Hamburg
Vogt-Kölln-Straße 30, D-22527 Hamburg
drewniok@informatik.uni-hamburg.de
http://kogs-www.informatik.uni-hamburg.de/staff/Drewniok_Christian.html

Commission III, Working Group 2

KEY WORDS: Vision, Automation, Extraction, Orientation, Registration, Urban Scene, Model-Based Detection, Invariants

ABSTRACT

We present an approach for automatic exterior orientation of aerial images which is based on the use of manhole covers as landmarks. The approach includes two main procedures: First, a landmark extraction scheme which enables us to automatically detect and precisely localize many of the manhole covers visible in an image. Second, an automatic matching procedure which robustly and efficiently matches constellations of detected landmarks with the correct landmarks from a cadastral database. By combining both methods we are able to automatically generate a large number of landmark correspondences per image, which allows for a reliable estimation of the exterior orientation parameters.

1. INTRODUCTION

The development of operational procedures for the automatic orientation of aerial images is a matter of topical interest in photogrammetric research (cf. Gülch (1995)). It requires the automation of each step in the orientation process. For absolute orientation this process includes:

1. Estimation of the interior orientation by detecting, precisely localizing, and identifying the fiducial camera marks, knowing the type of the camera used.
2. Detection and precise localization of landmark features within the image, assuming that these features correspond to known geodetic coordinates.
3. Identification of the features by determining correspondences between the landmarks detected in the image and existing landmarks from the observed scene.
4. Estimation of the exterior orientation by spatial resection based on the landmark correspondences.

For automatic interior orientation operational techniques have recently been worked out (Schickler, 1995), which make use of the existence of well-defined geometric models of the fiducial marks and profit from the well-structured appearance of these marks within the image (completely isolated features with high contrast). Opposed to this, in the automation of exterior orientation one has to deal with real scene objects and their complex appearance in aerial images. As a consequence, only very few approaches to automatic exterior orientation have been developed so far (see (Schickler, 1992) and (Vosselman & Haala, 1992)).

In this contribution we present an approach to automatic exterior orientation which is based on a specific type of circular landmarks. We suggest that manhole covers are well suited features to serve as landmarks for orientation of urban scenes. The advantages of using this kind of landmark are manifold: A great number of manhole covers can be found in urban environments, most of them being placed in the middle of a road; they are well distributed and located at the ground; geodetic coordinates (including altitudes) are available from the cadaster of the city's sewerage system; and, as we will show, they can be detected, localized with high precision, and associated with geodetic coordinates from the cadastral database, all in an automatic manner.

This contribution elaborates on the two major aspects in this context: First, how to detect landmarks of the considered type automatically and, second, how to match constellations of detected landmarks with the manhole positions from the cadastral database. Note that in the following the interior orientation of the images is assumed to be known.

2. MODEL-BASED DETECTION AND LOCALIZATION OF LANDMARKS

Our landmark extraction approach is based on a parametric model which explicitly describes the location, size, shape, and systematic intensity variations of depicted manhole covers. Minimizing the squared intensity error between the model and the image results in the best-fit model parameters and, most important, determines the landmark's position in the image with high sub-pixel precision. This can be shown for simulated as well as for real image data. A short description of the landmark extraction approach will be given below. More details can be found in (Drewniok & Rohr, 1995).

2.1. Analytic Description of Circular Landmarks

While the appearance of manhole covers varies from country to country we frequently find a specific type which consists of a bright disk surrounded by a dark concentric ring (see Figure 1, left). Since aerial images normally are recorded approximately parallel to the ground plane, images of these objects are circular. The idealized image intensities of a cross-section through a manhole cover of the considered type form a symmetric step function. When we also take into account that the intensities are blurred because of the band-limiting effect of the camera, we get a rounded shape as sketched in Figure 1 (right). This profile can approximately be described by three characteristics: h_{max}, h_{min}, and r_{min}, where h_{max}

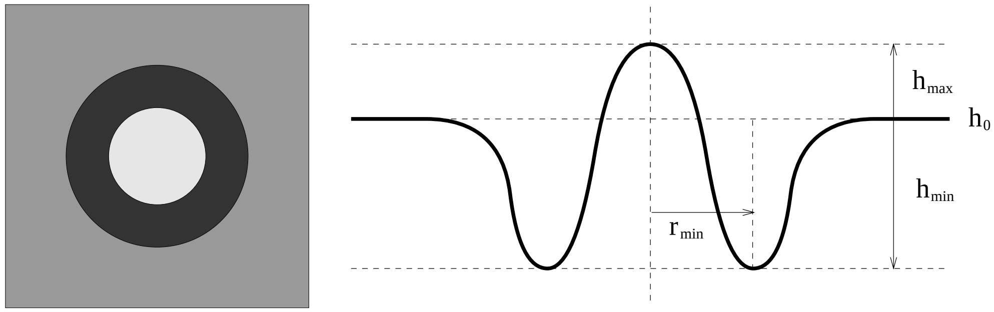

Figure 1: Ideal appearance of a manhole cover (left) and blurred cross-section intensities (right).

and h_{min} are the relative values of the function's maximum and minimum with respect to the background-level h_0; r_{min} denotes the distance of the minimum from the center. As suggested by Figure 1, we approximate the ideal intensity profile using an analytic model whose general shape corresponds to the second derivative of the 2D Gaussian. However, the shape of this function is controlled by only two parameters (amplitude and variance), while three parameters are needed for adequately describing the intensity profile of the landmark. Therefore, we represent the model by a modified function, which, on the one hand, well approximates the second derivative of a Gaussian, and, on the other hand, has three parameters describing its shape, namely a_1, a_2, and σ:

$$M(x,y) \;=\; a_0 + (a_1 + a_2 \cdot r^2) \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right) \qquad (1)$$

with $r^2 = (x - x_0)^2 + (y - y_0)^2$ and (x_0, y_0) being the image coordinates of the landmark center.

Given our model function M and the image intensities of an instance of a manhole cover, we minimize the error E between the image and the model. In our approach, E is defined by the sum of the squared differences between the image intensities I and the model M (which is a function of a_0, a_1, a_2, σ, x_0, and y_0) at some data points taken from a square window centered around the initial location estimate of the landmark. Since we are dealing with a non-linear model, we apply the iterative Levenberg-Marquardt method for minimizing the error function. This method requires to analytically calculate the first partial derivatives of M with respect to each of the parameters to be optimized. An example for a model fitting result is given in Figure 2.

2.2. The Landmark Extraction Scheme

In our landmark extraction scheme, we obtain an initial set of potential landmark positions by exploiting the normalized cross-correlation for all local intensity maxima in the image, using a landmark prototype template. The potential landmarks detected this way are submitted to the parameter optimization procedure described above which adapts the analytic model function to the intensities of the given landmark. The results of the model fitting are twofold: the set of optimally adapted parameters and the final approximation error. Both are checked in a subsequent verification step in order to decide whether the adapted model describes a valid landmark instance. In this way, we are able to suppress a large fraction of false detections and obtain a high-precision localization of the actual landmarks.

The initial parameter values required to set up the parameter optimization process are obtained from a small number of representative examples which have to be selected by the operator in a preceding training phase. The training results are also used to derive the thresholds applied in the verification step and to generate the prototype template used for landmark detection.

2.3. Localization Precision Obtained by Model Fitting

We investigated the localization precision obtained by model fitting for simulated landmark images which have been generated with known parameters. The physical landmark size was assumed to be 80 cm; the acquisition parameters and the pixel size were set to typical values (see below), giving a pixel resolution of 15 cm on the ground, i.e. a landmark is typically represented by 6 by 6 pixels (unblurred). Using our simulation technique we are able to statistically evaluate the localization precision with respect to variations in image blur, sampling effects, noise, perspective projection, regular shape distortions, and interaction effects with background structures (e.g. road markings). In a large number of random experiments we found that the localization error is well below a hundredth of a pixel (less than 1 mm on the ground) for noise-free images and less than a tenth of a pixel (about 1 cm on the ground) for images having a realistic amount of noise. Even in the presence of serious background distortions the localization precision is still in the lower sub-pixel range for most kinds of distortions.

2.4. Extraction Performance on Real Image Data

Experiments on real image data have been done on a number of data sets. We used color infrared photographs, which have been acquired with a Zeiss RMK A-30/23 camera from an altitude of 1500 m, giving an image scale of 1:5000. The photographs have been scanned at a resolution of 30 μm, resulting in a pixel size of 15 cm on the ground and an image size of about 7700 by 7700 pixels (covering a ground area of about 1.3 km^2).

A number of 400 to 500 manhole covers is visible in each of the images, including a significant fraction of landmarks which are seriously distorted by background structures, have very low contrast, or do not agree with the ideal landmark model at all. According to these effects, the extraction process typically yields a number of 100 to 200 detections. The percentage of false detections is in the range of 10% to 20% of the total number of detections, which is very low considering the high complexity of the analyzed scenes.

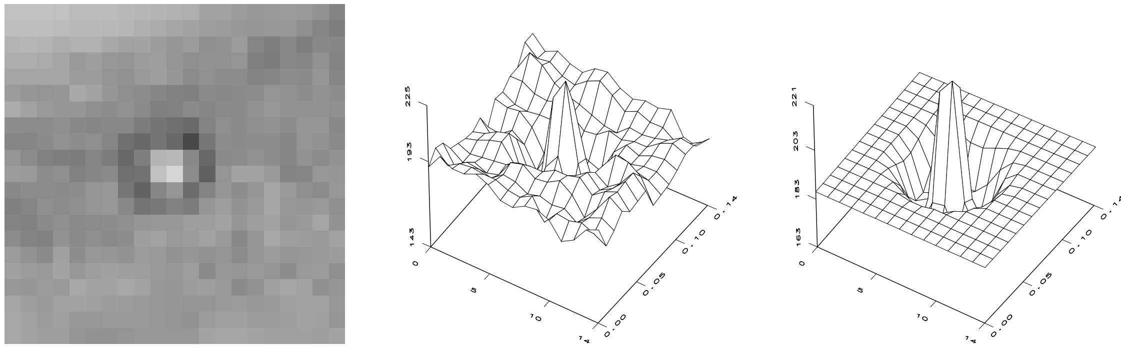

Figure 2: Manhole cover: intensities (left), intensity plot (center); adapted model (right).

Thus, in our experiments we are able to automatically extract 25% to 30% of all the landmarks visible in the images, while the number of false detections is reasonably low. Figure 3 illustrates a result obtained on a small image subpart.

3. MATCHING LANDMARK CONSTELLATIONS WITH A LANDMARK DATABASE

Given both, the positions of the landmarks detected in an image and a database containing the geodetic coordinates of all landmarks that are known to exist in the observed area, we now have to solve the problem of identification, i.e., to find the correct correspondences between the two sets. Finding these correspondences automatically provides us with the ground control information required to perform an estimation of the exterior orientation parameters automatically.

3.1. The Nature of the Matching Task

An important characteristic of the given matching task is that, in general, we cannot identify or distinguish individual manhole covers on the basis of their visual appearance. Therefore, a matching algorithm can only exploit geometric relations of sets of manhole covers. This implies that we cannot decide about the correctness of a single correspondence; rather, correctness is defined on a set of correspondences: To represent a correct total match, the set has to be geometrically consistent, i.e. there must be a perspective transformation which relates a sufficiently large fraction of the image landmarks with database landmarks resulting in sufficiently small residuals.

It is also important to note that we have to match large sets of landmarks. This demands an effective control of the search mechanism in order to prevent a combinatorical explosion. It is obvious that, for combinatorical reasons, we are not able to systematically try all possible sets of point-to-point matches. Instead, the matching process is based on characteristic *constellations* among the landmark sets whose specifity reduces the ambiguities in the candidate selection process. In the following, constellations of image landmarks will be termed *image constellation* while we denote constellations of database landmarks by the term *database constellation*.

One possibility would be the use of very specific constellations like subparts of the road network, knowing that this network is clearly reflected in the spatial arrangement of the manhole covers. However, we decided not to use complex types of constellations for two reasons: First, only a small subset of the landmarks included in the database is actually visible and detectable within the image, and second, the set of detections includes a significant fraction of landmarks which do not correspond with a database landmark (false detections as well as manhole covers which are located on private ground and, thus, are not included in the municipal register). As a consequence, there might be a very low chance that a complex pattern defined on the database landmarks can also be found within the set of detections. In order to deal with this kind of incompleteness and erroneousness we decided to use rather simple types of constellations (to be discussed below), thereby accepting that we are still faced with a lot of ambiguities in this case.

3.2. Matching Constellations by Hypothesize-and-Test

We overcome the problem of ambiguity by applying a hypothesize-and-test mechanism: First, we randomly select a set of image landmarks which make up a valid constellation. For this image constellation we determine all possibly matching database constellations. This is done by indexing through geometric invariants for reasons of efficiency. Given an image constellation and several candidate database constellations, we then *hypothesize* a match between the image constellation and one of the candidates. As indicated above, we can verify (*test*) such a hypothesis by estimating the orientation parameters from the landmark correspondences given by the constellation match and scoring the number of "hits". The number of hits is determined by transforming the coordinates of the image landmarks into world coordinates (which presumes that a digital elevation model is available) and doing a nearest-neighbor search among the database landmarks within a small search radius (e.g. 1.5 m in physical dimensions). When the number of hits obtained in this way is too small (i.e. only a small fraction of the detections can be "explained" in this way), we consider the next candidate from the indexed database constellations. If the verification fails for all candidates, we randomly select a new image constellation. The idea behind this method is that it will succeed with a large number of hits for perfect image constellations, while for a faulty image constellation it fails yielding only a small number of hits. By the word *perfect* we denote those image constellations which do not include any false detections or manhole covers not registered in the database. Knowing the size of the constellations used

Figure 3: Results of the landmark extraction scheme applied to a small sub-image which shows five manhole covers. On the left side, small circles indicate the potential landmarks detected by normalized cross-correlation. In the final result (right) the cross centers correspond to the estimated landmark centers while the radii of the circles correspond to the values of r_{min}. Note, that one of the visible landmarks has been rejected by the verification procedure due to its high approximation error.

as well as the average ratio between correct detections and erroneous ones (from our extraction experiments), we can specify the mean number of random trials required to select a perfect image constellation, which should give a successful total match. In summary, the algorithm proceeds as follows:

1. Prepare the database by constructing all database constellations according to the rules to be specified below.
2. Randomly select a set of image landmarks which makes up a valid image constellation.
3. Determine all possibly matching database constellations through geometric indexing.
4. Select an unused candidate from the indexed database constellations.
5. Estimate the orientation parameters based on the given landmark correspondences and score the number of hits.
6. When the number of hits is higher than a given threshold, stop with success; otherwise, when there is still an unused candidate, continue with step 4; otherwise continue with step 2.

Note that in case of a successful total match, the desired result—i.e., the exterior orientation—is already available from the verification step which determines the orientation parameters based on a maximum number of landmark matches. Since we can expect to yield a high number of correspondences per image (50–100), the estimate is highly reliable. Thus, based on a reliable estimate it is possible to recognize landmarks with significantly distorted coordinates by looking for outliers in the residuals. Such an analysis enables us to finally enhance the result by excluding detected outliers as well as to examine the accuracy of the coordinates included in the landmark database.

3.3. The Type of Constellations Used

We now have to define the type of constellations to be used in our approach. The matching algorithm sketched above implies a number of criteria to be considered in this context:

- It should be possible—with respect to time and storage capacities—to precompute all valid constellations among the database landmarks.
- There should be a high probability that constellations of the desired type can be found among the landmarks detected in the image. This excludes complex constellations which are specific but also sensitive to missing or additional features.
- There must be an efficient method to access all possibly matching database constellations for a given image constellation.
- To limit the computational effort required to estimate the orientation parameters and to transform the coordinates, the number of candidate database constellations associated with a given image constellation should be small.
- The probability for randomly selecting a perfect image constellation should be high.

Trying to find a trade-off between specifity and robustness, we considered two types of constellations: triples and five-tuples. The most serious argument for using small-sized n-tuples is given by the last criterion from the list above: The probability for randomly selecting a perfect n-tuple decreases exponentially with increasing n; Table 1 reveals the consequences of this relationship. The experiments reported in Section 2.4 have shown that the set of detections typically includes a fraction of false detections in the range of 10% to 20% (see

tuple size	#trials for an error rate of			
n	20%	30%	40%	50%
3	2.0	2.9	4.6	8.0
4	2.4	4.2	7.7	16.0
5	3.1	5.9	12.9	32.0
10	9.3	35.4	165.4	1024.0

Table 1: Mean number of trials required to randomly select a perfect n-tuple of detected landmarks.

above). Moreover, about 20% of the detections typically represent manhole covers which are visible in the image but which are not included in the cadastral database, for several reasons. Thus, we have to expect that only about 60% of all detections actually correspond with a cadaster landmark (i.e. the "error" rate here is about 40%). This clearly favors the use of small-sized constellations in order to keep the number of random trials as small as possible.

To enable the precomputation of all database constellations we have to restrict the resulting combinatorics. This is done by constraining the pairwise distances of the constellation landmarks. Thus, for a given database landmark we only have to consider its neighbors located within a certain minimum and maximum distance in order to construct all possible constellations. Searching for the neighbors can be supported by using a spatial index provided by a spatial database. Additionally, for a triple we demand that the three altitudes of the "triangle" represented by the three landmarks also exceed the minimum distance threshold. The purpose of this rule is to exclude nearly collinear tuples which would lead to less reliable estimates for the orientation.

Since we deal with nearly vertical imaging conditions and with more or less flat urban areas, in a first approximation distances can be assumed to be "invariant" up to a common scaling factor. Thus, the minimum and maximum distance thresholds can be transformed to pixel distances knowing the approximate value of the image scale and the pixel size. Using the scaled distance thresholds we can construct image constellations in the same way as described for the database constellations. The distance thresholds should be small enough to effectively restrict the total number of valid constellations in the precomputation phase. They should be large enough to ensure that we can find a sufficient number of detections within the resulting search area in the image, giving a sufficiently high probability for constructing perfect image constellations. In general, we can assume that the deviations from vertical geometry are sufficiently small to ensure that for a valid database constellation there exists a valid image constellation—provided that the involved landmarks are actually detected in the image—and vice versa. However, we have to accept that we will miss some constellation matches due to the uncertainties resulting from deviations from vertical imaging, from neglecting elevation variations, and from the limited accuracy of the cadastral coordinates.

3.4. Candidate Selection Through Indexing

As mentioned above, our approach includes an efficient candidate selection mechanism which is based on indexing through certain geometric properties of the constellations. While, in principle, this technique is applicable to both, triples as well as five-tuples, there are some arguments which favor the use of triples here. In order to justify our decision to use triples in our final implementation, we discuss these arguments below.

It is well known that for a set of correct correspondences the positions in the image and in the world are related by a projective transformation; moreover, we know that there exist certain geometric properties of point constellations which are invariant under perspective projection, e.g. the cross-ratio of four collinear or five coplanar points. Thus, using five-tuples (and assuming that the 3D-coordinates of the landmarks do not deviate from a plane too much) would suggest to use their cross-ratio as indexing key. In this case, we would precompute the cross-ratio for all database constellations and make according entries in an index table. In the matching phase we would compute the cross-ratio for each given image constellation and use it as an index to this table (see (Meer et al., 1993) and (Lenz & Meer, 1994) for an instructive discussion on the use of projective invariants for matching). However, due to the uncertainties mentioned above and due to geometric similarities that might occur among the individual constellations, we have to accept that there is no unique relationship between image constellations and database constellations. While, in our application, this holds for most kinds of geometric properties, the use of the cross-ratio entails some specific problems: First, requiring a number of five landmarks per constellation increases the mean number of random trials required in the matching process (see above). Second, using the full degrees of freedom provided by a perspective projection does not reflect the strong restrictedness of the nearly vertical imaging geometry. As a consequence, the matching algorithm will be faced with some additional ambiguities which could be excluded under the assumption of almost affine conditions. And third, the uncertainties mentioned above will lead to non-ideal correspondences of the cross-ratio values to be compared. While this is inevitable, we have to note that it is hard to assess distances of ratios due to their highly non-linear nature.

Therefore, we prefer to use triples instead of five-tuples in our matching approach and decided to implement an indexing method based on their ordered set of landmark distances. Using only three landmarks gives a low number of random trials required; analyzing their (scaled) distances does reflect the restricted imaging conditions quite well; and it is also more obvious how to evaluate Euclidean distances between points than distances of cross-ratios. The implemented indexing mechanism proceeds as follows: To get a unique representation for a given triple we first put the three coordinates in a counter-clockwise circular order according to their spatial arrangement and compute the distances between each point and its successor (in the case of image landmarks, the pixel distances are transformed into world distances by scaling with the approximately known scaling factor). Finally, we create a linear list of the three distances following the counter-clockwise order of the points and starting with the largest distance value. This gives a uniquely defined distance vector carrying its largest element in the first place. For the database triples we use this vector to create a three-level index table. Thus, given a triple of image landmarks we compute its associated distance vector and use this for indexing the candidate table which requires three table access operations.

3.5. Verification by Iterative Registration

For a given set of three landmark matches we are able to estimate the image orientation parameters by spatial resection. This technique is used in our matching algorithm for verifying a hypothesized constellation match. However, knowing that for small sets of correspondences the estimation of the orientation might be unreliable, we can enhance our verification criterion by applying an iterative registration method: For a perfect image constellation the initial estimate of the orientation will probably be good enough to instantiate a number of additional correspondences by projection and nearest-neighbor search. These additional correspondences, when included in the parameter estimation process, allow for a more reliable estimate, which again results in an increased number of correspondences. So, by repeatedly applying this combined step of parameter estimation, projection, and nearest-neighbor search we obtain a successively increasing number of landmark correspondences. The iteration process stops when the set of correspondences remains unchanged during the current step. We can assume that the iterative mechanism yields a large number of correspondences covering a significant amount of all the detected landmarks when a perfect image constellation is used and associated with the correct database constellation. In the other cases, the iteration process stops quickly, yielding only a small number of correspondences. Hence, in order to decide on the success of the matching process we can simply exploit the final number of correspondences.

In order to limit the computational effort claimed by the iterative registration we only start the iteration process for those constellation matches which yield a number of at least five correspondences using the initial orientation estimate. We will see below that the number of constellation matches to be submitted to the iterative procedure can be reduced considerably by applying this heuristic.

3.6. Experimental Results

Our approach for matching landmark constellations was implemented and tested on real imagery, using automatically detected landmarks and geodetic data from a cadastral database covering the observed scene and the surrounding area. Below we will discuss results obtained for an image which shows the central area of Harburg (southern part of the city of Hamburg) and has been acquired in July 1987 at about 4 p.m. (see Section 2.4 for technical parameters). Our geodetic landmark database includes a total number of 877 manhole covers from an area covering about 2 km^2 (recall that the image covers about 1.3 km^2). About 600 manhole covers are located within the area covered by the image, while about 300 covers are actually visible in the image (including those which are distorted, which have low contrast, etc.). We found in a number of tests that for images with a scale of 1:5000 we obtain satisfying matching results with distance thresholds of 50 m (minimum distance) and 150 m (maximum distance). Using these thresholds, from the 877 database landmarks the algorithm constructed a number of 74382 triples which fulfill the pairwise distance criterion as well as the triangle altitude criterion. The resulting distance vectors were used to create a three-level index table. Here, we used table bins in a discretization of 5 m to take into account the uncertainties which might affect the distance measurements among image landmarks.

For the Harburg image the landmark extraction scheme obtains a number of 131 detections, including 23 false detections and 25 manhole covers which are not included in the cadastral database. Thus, for the matching process the error rate mounts up to about 37%; in the best case, the matching approach can yield a number of 83 correct correspondences. Among the 131 detected landmarks the algorithm constructs 688 valid triples, including 213 perfect image constellations. We see that the ratio of 688 to 213 approximately agrees with the value predicted by Table 1 for an error rate in the range of 30% to 40%. Accordingly, in a random selection mode the mean number of trials required to select a perfect image constellation is in the range of 3 to 4, or, in other words, there is a chance of about 30% to randomly select (or construct) a perfect image constellation.

In the experiments, the acceptance threshold defining a successful total match was set to 50% of the number of detections (i.e., a minimum of 65 correspondences is required). To enable the projection of image coordinates into world coordinates, we created a digital elevation model from the 3D-coordinates of the landmarks contained in the cadastral database using Delaunay-triangulation. Note that this triangulation at the same time provides the Voronoi-diagram of the database landmarks which we used to implement the nearest-neighbor search in a very efficient way. The search radius was set to 1.5 m. To investigate the behavior of the matching algorithm we applied it to each of the 688 triples. The results can be summarized as follows:

- The mean number of candidates accessed by our indexing method is about 180 (to take into account uncertainties of the distance measurements each distance value is used to access two neighboring bins of the candidate table instead of exactly one, giving a mean tolerance of about ± 5 m).
- By doing the initial orientation estimation and computing the number of hits, the candidate set is reduced to 1 or 2 potential constellation matches which yield the minimum number of five initial landmark correspondences (1.55 on the average).
- Only 40% of all the image constellations pass this initial test (i.e. there is at least one potential constellation match passing the test): 87% (186) of the perfect image constellations and only 19% (92) of the erroneous ones. Thus, the iterative registration procedure has to be applied to a very small number of potential constellation matches only.
- From the 92 erroneous image triples passing the initial test there are only two which yield a final number of correspondences above 10 when performing the iterative registration method. One of the two yields 18 correspondences which is uncritically low with respect to the acceptance threshold (65). The second one leads to a successful total match yielding the maximum number of 83 correct correspondences (probably, the initial constellation match was not far from perfect). So there is no falsely accepted total match and for correctly rejected total matches the final number of hits is much lower than the acceptance threshold.

- From the 186 perfect image triples 149 succeed when submitted to the iterative registration procedure. There are 37 constellations which yield a final number of hits in the range of 10 to 45 which is again significantly lower than the acceptance threshold. Note, that a set of 45 correspondences implies that there are at least 38 correct matches which were missed due to too large residuals. Thus, the resulting orientation estimate cannot be satisfying and, consequently, rejecting a total match which does not yield the maximum number of 83 correspondences is appropriate and justified.

- The number of 149 successfully matched image triples represents 70% of all perfect constellations and 22% of the total number of valid image triples. Thus, we can conclude that on average we have to perform 4.5 random trials until the matching process signals success. All inappropriate constellation matches are guaranteed to be rejected, while the choice of the acceptance threshold is highly uncritical.

4. SUMMARY AND CONCLUSION

We presented an approach for automatic exterior orientation of aerial images which is based on the use of manhole covers as landmarks. For this purpose we developed a landmark extraction scheme which enables us to automatically detect a quarter up to a third of all manhole covers visible within an image while the error rate is reasonably low. Using the positions of the detected landmarks for estimating the orientation parameters requires to identify each landmark, i.e., to find the correct correspondence with a landmark from the cadastral database. This database includes the geodetic coordinates of all manhole covers that exist within the observed area. Thus, we have to match landmark constellations, where the set of image landmarks is incomplete and distorted (undetectable covers and false detections, resp.). We introduced a matching approach which is based on matches of triples of landmarks. In a randomized selection process a triple match is hypothesized; the match is verified by estimating the orientation parameters and looking for additional correspondences, both combined in an iterative process. Candidate selection is efficiently implemented using indexing through geometric invariants. In our experiments the matching approach proofed to be successful: It robustly yields the correct total match. Based on a large number of correspondences, the final match provides a reliable estimate of the exterior orientation parameters.

We conclude that by combining both, a robust and precise landmark extraction scheme with a reliable matching mechanism, we are able to provide an effective approach to automatic exterior orientation of aerial images. Our approach offers the opportunity to automate the orientation process in urban environments, presuming that the approximate image scale and the approximate location of acquisition can be derived from the flight plan. It requires that the geodetic coordinates of manhole covers are available from cadastral databases in digital format. This might not be the typical case today, but triggered by the transition from analog to digital photogrammetry and by the rapidly spreading use of GIS-technology, there is a strong trend towards digitalization in many cadastral tasks.

SOURCES OF DATA

Aerial photographs and map material used in our experiments have been kindly provided by the Vermessungsamt Hamburg. For illustrations we used the image FIR 87-20-828 acquired by the Gesellschaft für technische Photogrammetrie m.b.H, Hamburg. Reproduction is done with the permission of the Vermessungsamt Hamburg. Digital coordinates extracted from the cadaster of the sewerage system of the city of Hamburg have been kindly provided by the Hamburger Stadtentwässerung, Anstalt des Öffentlichen Rechts, Hamburg.

ACKNOWLEDGMENTS

We thank Rainer Sprengel for constructive suggestions with respect to the landmark model, Joachim Rieger for suggesting the use of invariants, and Carsten Schröder for support in the implementation and for making valuable contributions to the revision of this manuscript. For providing the data material we thank Mr. Rohwedder (Vermessungsamt), Mr. Sachweh (Vermessungsamt), and Mr. Gleim (Stadtentwässerung). Parts of the work reported here have been done in the LUKAS-project carried out at the AI-Laboratory at the Computer Science Department, University of Hamburg, in cooperation with the Ingenieurbüro Basedow and Tornow GmbH, Hamburg.

REFERENCES

Drewniok, C. and K. Rohr, 1995. High-Precision Localization of Circular Landmarks in Aerial Images. In Sagerer, G., S. Posch, and F. Kummert, editors, *Mustererkennung 1995 – 17. DAGM-Symposium*, Bielefeld, Germany, 13.–15. September 1995, Informatik aktuell, pages 594–601. Springer-Verlag, Berlin – Heidelberg – New York, 1995.

Gülch, E., 1995. From Control Points to Control Structures for Absolute Orientation and Aerial Triangulation in Digital Photogrammetry. *Zeitschrift f. Photogr. u. Fernerkundung* **63** (3), 130–136, May 1995.

Lenz, R. and P. Meer, 1994. Point Configuration Invariants Under Simultaneous Projective and Permutation Transformations. *Pattern Recognition* **27** (11), 1523–1532, 1994.

Meer, P., S. Ramakrishna, and R. Lenz, 1993. Correspondence of Coplanar Features Through P^2-Invariant Representations. In Mundy, J.L., A. Zisserman, and D.Forsyth, editors, *Applications of Invariance in Computer Vision*, Ponta Delgada, Azores, Portugal, October 1993, volume 825 of *Lecture Notes in Computer Science*, pages 473–492. Springer-Verlag, Berlin – Heidelberg – New York, 1994.

Schickler, W., 1992. Feature Matching for Outer Orientation of Single Images Using 3-D Wireframe Controlpoints. In *Proc. XVII ISPRS Congress*, Washington, D.C., 1992, volume XXIX, part B3 of *International Archives of Photogrammetry and Remote Sensing*, pages 591–598. American Society for Photogrammetry and Remote Sensing, Bethesda, MD, 1992.

Schickler, W., 1995. Ein operationelles Verfahren zur automatischen inneren Orientierung von Luftbildern. *Zeitschrift f. Photogr. u. Fernerkundung* **63** (3), 115–122, May 1995.

Vosselman, G. and N. Haala, 1992. Erkennung topographischer Paßpunkte durch relationale Zuordnung. *Zeitschrift f. Photogr. u. Fernerkundung* **60** (6), 170–176, Nov. 1992.