High-Precision Localization of
Circular Landmarks in Aerial Images

Christian Drewniok and Karl Rohr[1]

Fachbereich Informatik, Universität Hamburg
Vogt-Kölln-Straße 30, 22527 Hamburg

Abstract. The reliability and accuracy of point-based image registration strongly depends on the selection of suitable landmarks and on the precision of localizing theses landmarks in images. In this contribution, we consider the problem of landmark extraction for the purpose of aerial image registration. We suggest to use a specific type of circular landmarks and introduce a model-based approach for localizing these features with high subpixel precision. The approach has been tested on synthetic as well as on real image data.

1 Introduction

Accurate registration of aerial images is essential to any kind of their photogrammetric exploitation. Knowing the exact positions of control points on the ground and in the image enables to reconstruct the imaging geometry. Often, ground control points are premarked to aid in their detection and localization in the image (artificial landmarks). If the photo mission has not been prepared this way, selecting suitable points in large-scale photographs (e.g. 1:5000) is difficult: features like road-intersections in general cannot be used, since they do not bear a unique location in highly resolved images; corners of buildings are usually not visible due to occlusion; roof structures are displaced due to elevation and must be excluded when heights are unknown.

In this contribution, we suggest that manhole covers placed in the middle of streets are well suited features which can serve as landmarks for registration of urban scenes. The advantages are threefold: a great number of manhole covers can be found in urban environments; they are well distributed and located at the ground plane; geographic data is available; and, as will be shown below, they can be automatically detected and localized with high precision in aerial images.

Our work is based on a parametric model which explicitly decribes the systematic intensity variations of depicted manhole covers. By fitting this model directly to the image intensities the landmarks can be localized to high subpixel precision. Also, it is possible to verify the fitted model, either on the basis of the estimated parameters or by exploiting the approximation error. Moreover, the detection of the landmarks can greatly be supported by a prototype model which in our case is determined using a simple learning scheme.

Previous work on the extraction of circular landmarks has been concentrated on indirect approaches, e.g. fitting circles to grey-value edges. Other approaches do not exploit an explicit model (e.g. [2]). Fitting approaches comparable to ours have been developed for extracting low-level image features, namely edges [3] and corners [4]. However, the intensity models of these approaches only describe parts of a depicted object, while our model represents the systematic intensity variations of the entire object.

2 Analytic description of circular landmarks

We frequently find a specific type of manhole covers which consists of a bright disk surrounded by a dark concentric ring (see Fig. 1, left). Since aerial images normally are recorded 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. Also considering that the intensities are blurred because of the band-limiting effect of the camera gives a rounded shape as sketched in Fig. 1. This profile can approximately be described by 3 shape characteristics: h_{max}, h_{min}, and r_{min}, where h_{max} 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 position.
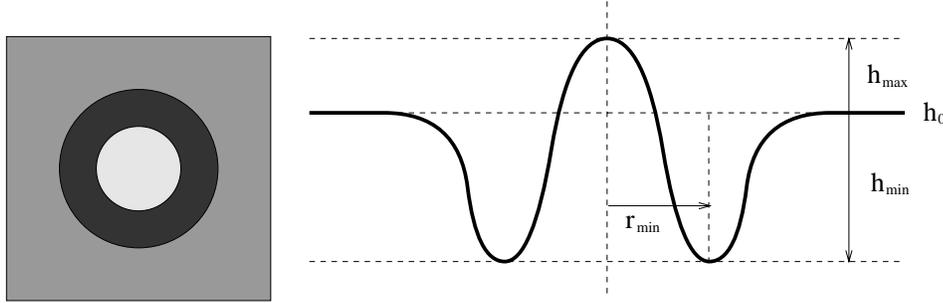

Figure 1: Ideal appearance of a manhole cover (left) and blurred cross-section intensities (right).

Figure 1 (right) suggests to use 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 we need three parameters for describing the shape of the landmark. We therefore represent the model by an adapted version of the difference of two Gaussians, which, on the one hand, well approximates the second derivative of a Gaussian, and, on the other hand, has three shape parameters, namely a_1, a_2, and σ:

$$M(r) = a_0 + (a_1 + a_2 \cdot r^2) \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right) . \tag{1}$$

While simply $h_0 = a_0$ and $h_{max} = a_1$, the following relationships hold for h_{min} and r_{min}:

$$h_{min} = 2\sigma^2 a_2 \cdot \exp\left(-\frac{2\sigma^2 a_2 - a_1}{2\sigma^2 a_2}\right) \qquad r_{min} = \sqrt{2\sigma^2 - \frac{a_1}{a_2}} . \tag{2}$$

Assuming $h_{min} < 0$ and $h_{min} < h_{max}$, (which is a weak restriction in our application) we can show that these equations can efficiently be solved for a_2 and σ [1]. Hence, we are able to specify initial values of the model parameters from an estimate of the landmark characteristics.

3   A computational approach to landmark extraction

A general scheme for landmark extraction is sketched in Fig. 2. Its central component is a parameter optimization procedure which adapts the analytic model function to the intensities of a given landmark candidate. The output of this procedure is twofold: a set of adapted parameters and the approximation error. Both can be used in a subsequent verification step to check whether the adapted model describes a valid landmark instance. The model fitting procedure

has to be supplied with appropriate values for the initial parameter settings and the size of the observation window. As we will see, it is useful to determine the initial parameter values from a number of representative landmark examples. Positions of landmark candidates can be specified interactively or can automatically be detected using a template matching approach.
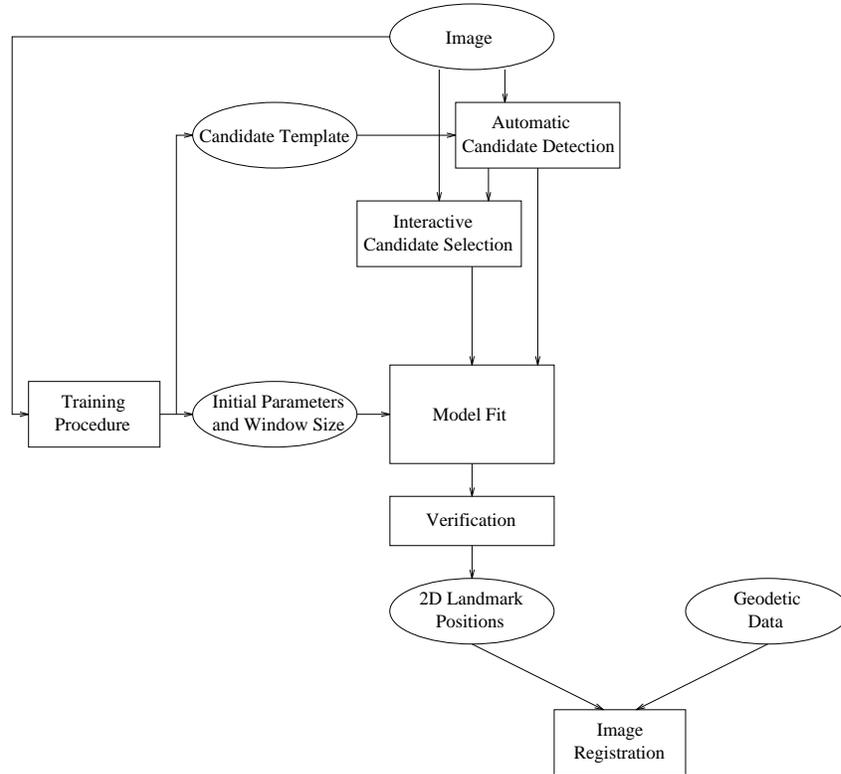
Figure 2: A general scheme for landmark extraction.

3.1 The fitting procedure

We minimize the squared error between the image intensities and the parametrized model. This results in the best-fit parameters, telling us shape, size, and subpixel location of the landmark. As we are dealing with a non-linear model, we use the iterative Levenberg-Marquardt method for minimizing the error function. The error is measured within a square window which is centered around the initial location estimate of the landmark. The width of the window is adjusted to the initial estimates of the model parameters in such a way that the absolute value of M (relative to h_0) falls below a certain fraction ($\frac{1}{100}$) of its largest value outside the window.

3.2 Learning initial parameters from examples

In order to obtain fast convergence and robust results, good initial parameter values are important. The initial landmark location is given by the position of the candidate, while initial values for the other model parameters are obtained from estimates of the landmark characteristics h_0, h_{max},

h_{min}, and r_{min}. The landmark characteristics itself are estimated on the basis of simple image measurements within a small subwindow around the landmark. The estimated landmark characteristics are transformed into values for the model parameters (using the relationships presented above) which are used to initialize the minimization procedure.

Although one could try to obtain initial parameters from image measurements for all given landmark candidates, we suggest not to do so. Since the landmarks are imaged with low resolution, we cannot expect simple image processing techniques to robustly estimate the landmark characteristics of each—possibly distorted—instance. We therefore prefer to apply this technique only to some well-structured, representative examples. This is done in a learning phase which precedes the actual landmark extraction process (cf. Fig. 2). The set of parameters resulting from taking the mean parameter vector of the learning examples is then used in the subsequent extraction phase for initialization of the minimization procedure.

3.3 Automatic candidate detection

For an efficient algorithm it is important to exclude false candidates as far as possible. This can be achieved by applying a number of selection criteria before the model fit is done. First, we can define the set of possible landmark positions by the set of local intensity maxima: each landmark will have a local intensity maximum within its bright center region and, thus, near its real center position. The second selection mechanism is based on template matching and proofed to be very effective. By using the learned landmark characteristics we can generate a prototype landmark template. This template can be used to filter the image and to reduce the set of candidates to those maximum positions which give high response in the filter output. To further reduce the number of candidates we finally exploit the initial fitting error. Each landmark candidate is tested against the initial landmark model (resulting from the learned parameters) without adapting any parameters. A candidate's initial fitting error has to be less than a certain threshold. Then, the model fitting procedure is applied to the candidates which passed all the three tests.

3.4 Final verification

The results obtained by minimization are submitted to a final verification applying five criteria:

1. The fitting error has to be less than a certain threshold.
2. h_0 has to lie within the intensity range.
3. h_{min} as well as $h_{min} - h_{max}$ have to be negative. Their absolute values have to lie within the intensity range and have to be higher than a specified minimum contrast.
4. r_{min} has to lie within $\pm 2.5\sigma_{r_{min}}$ around the mean of r_{min} observed for the learning examples.
5. The position may deviate from the inital candidate position by at most one pixel.

4 Experimental results

We fully implemented the landmark extraction scheme described above. In our experiments, we first investigated the localization precision obtained by applying our model fitting approach to simulated landmark images which have been generated with known parameters. The landmark size was assumed to be 80 cm; we further assumed typical acquisition parameters (see below) leading to 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, and some regular shape distortions. 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.

In the following we describe experimental results which demonstrate the overall performance of our landmark extraction scheme. These experiments are done on real aerial images typical for photogrammetric applications. We used (vertical) photographs with an image scale of about 1:5000. The photographs have been scanned at a resolution of 30 μm, so that one pixel approximately corresponds to a ground area of 15 by 15 cm. The results were obtained for the 600 by 600 pixel subscene shown in Fig. 3. Ten manhole covers are visible (see Fig. 3). One of these does not agree with our model (landmark `a'): its homogeneous background is darker than the landmark's ring, which cannot be represented by our model function.

Results on the test scene

The scene covers 360 000 pixels and yields 41 294 local maxima using a 3 by 3 neighborhood. The subsequent extraction results depend on the choice of the learning examples. All the nine landmarks have been tested on their suitability for the learning scheme. The scheme worked very well except for landmark 6 which required several trials to select a suitable subimage around the landmark. To illustrate the landmark extraction process we present in Fig. 5 the results obtained during the sequence of extraction steps when applied to a subpart of the test scene. For landmark 7 Fig. 4 shows the original intensity structure in comparison to the fitted model.

The results obtained by the extraction scheme for the complete test scene are summarized in Table 1. We find for each tested set of learning examples the number of candidates resulting from template matching (TM), the number of candidates remaining after exploiting the initial fitting error (IE), the number of hits, false negatives, and false positives, and the result of the final verification for each individual landmark. Obviously, there are two landmarks (1 and 6) which are very unstable to detect. The other seven landmarks are successfully detected independent from the choice of the learning set. However, for learning sets having a low standard deviation of r_{min} some of the landmarks failed during verification due to the r_{min}-criterion which is not surprising. In order to extract a wider range of differently sized landmarks the learning set has to represent this desired variability. By choosing a representative set of learning examples (e.g., 5, 7, 9) we are able to extract 8 of the 9 landmarks with only a few false positives. The majority of false candidates is refused by the final verification procedure (Sect. 3.4). In most cases (90%) more than two verification criteria became effective. This indicates robustness of the verification procedure and lessens the need for critical tuning the thresholds.

We have also analyzed the variance of the adapted position with respect to the learning set (see Table 2). We can see that the landmark position (x_0, y_0) estimated by the model fit is very stable (presuming that the learning set allows for the detection of a given landmark). The standard deviation of the x- and y-coordinates measured over all learning sets is well below a tenth of a pixel—in most cases it is better than a hundredth of a pixel. Another important aspect is the question of how landmark localization depends on the initial position estimate. Therefore, we compared the model fitting results for the stable landmarks while varying the initial position in a 3 by 3 pixel neighborhood around the actual candidate position (using landmarks 5, 7, and 9 as learning set). Only two landmarks result in detection failures for specific offsets. The standard deviation of the adapted locations is again well below a tenth of a pixel.
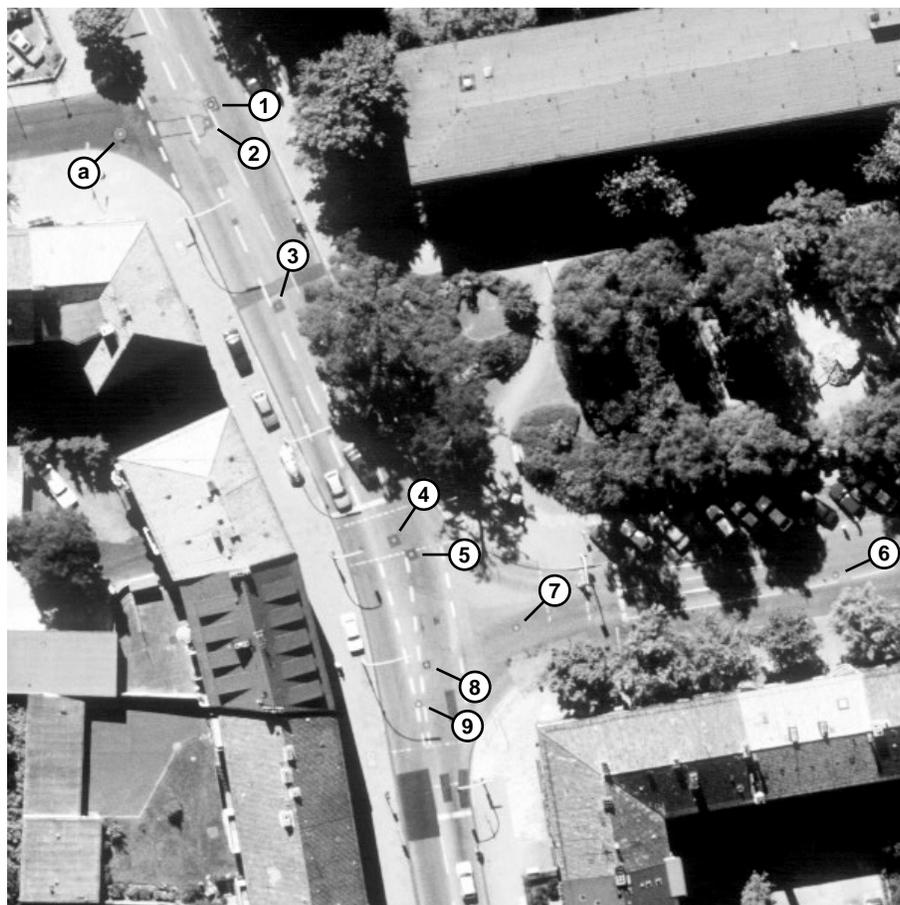
5

Figure 3: Scene used in the experiments. Ten landmarks are visible (markings 1...9 and `a').
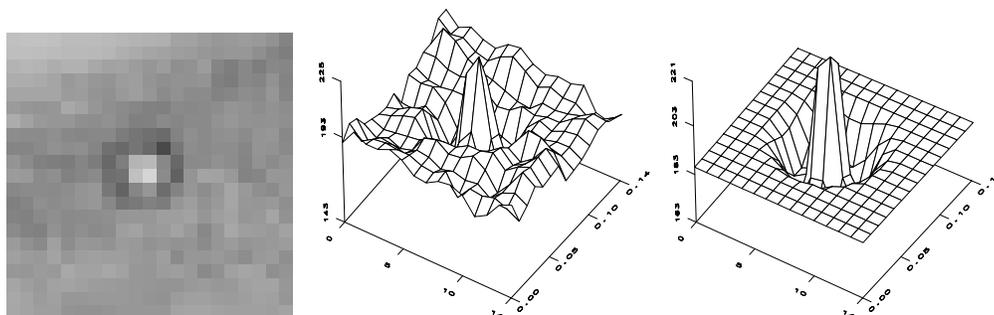

Figure 4: Intensities of landmark 7 (left, center) and of the fitted model function (right).
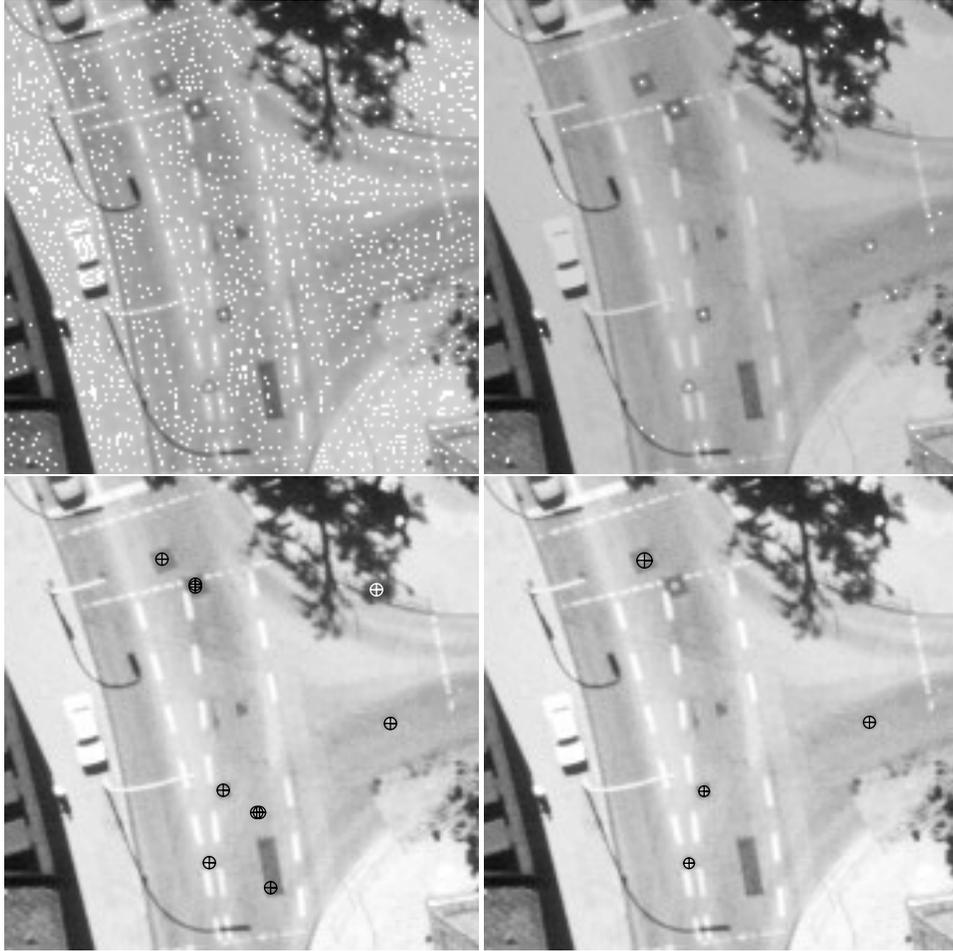
Figure 5: Results on a 170 by 170 pixels subimage using landmarks 5, 7, and 9 for learning: 2202 local maxima (top, left); 75 candidates remaining after template matching (top, right); 30 candidates with low initial error (bottom, left); 5 accepted landmarks (bottom, right).

5 Summary

We introduced a model fitting approach for detection and high-precision localization of circular landmarks which can serve for the registration of aerial images. By combining an effective candidate detection mechanism with our model fitting procedure we are able to robustly detect most of the landmarks visible in a complex scene automatically, yielding only few false responses. Both, the prototype template used for candidate detection as well as the initial model parameters used for the least-squares fit are determined from a few landmark examples by applying a simple learning scheme. Our approach has been shown to be efficient and robust. The localization precision was found to be well below a tenth of a pixel irrespective of typical variations in image blur, sampling, noise, or perspective distortions.

Table 1: Results of the landmark extraction scheme applied to the test image. `TM' stands for template matching; `IE' stands for initial error exploitation; + denotes success, - denotes failure, and r denotes failure only due to deviation of r_{min}.

landmark	# candidates		# hits	# false	# false	verification results								
examples	after TM	after IE		neg.	pos.	1	2	3	4	5	6	7	8	9
4,5	1227	78	4	5	2	-	+	+	+	+	-	r	r	r
4,7	1296	98	8	1	1	-	+	+	+	+	+	+	+	+
4,8	1083	50	7	2	0	-	+	+	+	+	-	+	+	+
4,9	1332	74	8	1	1	-	+	+	+	+	+	+	+	+
5,7	1826	201	6	3	2	-	+	+	r	+	r	+	+	+
5,8	1596	128	8	1	2	-	+	+	+	+	+	+	+	+
5,9	1827	154	8	1	3	-	+	+	+	+	+	+	+	+
7,8	1817	164	5	4	1	+	r	+	r	r	-	+	+	+
7,9	2183	209	4	5	2	+	r	r	r	r	-	+	+	+
8,9	1864	139	2	7	0	r	r	r	r	r	-	r	+	+
5,7,9	1548	108	8	1	2	+	+	+	+	+	-	+	+	+
4,5,7,8,9	1651	141	8	1	3	-	+	+	+	+	+	+	+	+

Table 2: Mean parameters of three of the stable landmarks when varying the set of learning examples (all pairs out of landmarks 4, 5, 7, 8, 9, plus combination 5, 7, 9, and the complete set).

landmark		h_0	h_{max}	h_{min}	r_{min}	error	x_0	y_0
5	μ	194.62	20.80	-68.77	2.52	19.11	268.295	362.787
	σ	3.81	5.78	3.67	0.03	0.04	0.012	0.001
7	μ	186.55	50.44	-23.47	2.15	5.78	337.577	411.579
	σ	0.25	0.48	0.28	0.00	0.06	0.001	0.002
9	μ	213.70	42.39	-32.26	1.92	12.23	272.832	462.215
	σ	0.50	0.87	0.54	0.01	0.26	0.001	0.004

Acknowledgments

References

[1] C. Drewniok. *Einsatz parametrischer Modelle in der Luftbildauswertung*. Dissertation, Fachbereich Informatik, Universität Hamburg, 1995. In Vorbereitung.

[2] W. Förstner and E. Gülch. A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Circular Features. In *Proceedings of the Intercommission Conference on Fast Processing of Photogrammetric Data*, Interlaken, 1987, pages 281–305.

[3] V.S. Nalwa and T.O. Binford. On Detecting Edges. *IEEE PAMI* **8** (6), 699–714, 1986.

[4] K. Rohr. Recognizing Corners by Fitting Parametric Models. *Int. J. Comp. Vision* **9** (3), 213-230, 1992.