Efficient Video Mosaicking by Multiple Loop Closing

Jochen Meidow

Fraunhofer Institute of Optronics, System Technologies and Image Exploitation
IOSB, 76275 Ettlingen, Germany
jochen.meidow@iosb.fraunhofer.de

Abstract. The rapid generation of aerial mosaics is an important task for change detection, e.g. in the context of disaster management or surveillance. Unmanned aerial vehicles equipped with a single camera offer the possibility to solve this task with moderate efforts. Unfortunately, the accumulation of tracking errors leads to a drift in the alignment of images which has to be compensated by loop closing for instance. We propose a novel approach for constructing large, consistent and undistorted mosaics by aligning video images of planar scenes. The approach allows the simultaneous closing of multiple loops possibly resulting from the camera path in a batch process. The choice of the adjustment model leads to statistical rigorous solutions while the used minimal representations for the involved homographies and the exploitation of the natural image order enable very efficient computations. The approach will be empirically evaluated with the help of synthetic data and its feasibility will be demonstrated with real data sets.

Keywords: image alignment, mosaic, loop closing, homography, exponential representation, parameter estimation

1 Introduction

1.1 Motivation

Up-to-date aerial image mosaics reveal valuable information for various applications such as change detection for disaster management or surveillance. Small Unmanned Aerial Vehicles (UAV) are convenient platforms to accomplish this task with moderate efforts and costs. Because of their limited cargo bay and payload usually only a single camera can be used leading to monocular imagery. Furthermore, the accuracy of navigation solutions from deployable inertial measurement units is usually too low to rely on it. Then again assuming planar scenes and chaining consecutive image pairs of the video stream suffers from error accumulation due to systematic errors and the uncertainty of feature tracking or matching. This becomes evident when a loop is detected — the arising drift appears in discrepancies at the joints. Therefore, strategies are needed to update the mosaic by detecting loops and adjusting the involved mappings.

1.2 Related Work

In Szeliski (2006) a comprehensive survey of mapping models is given together with a discussion on pros and cons of direct and feature-based image alignment, including global registration via bundle adjustment. These stitching techniques do not address the special nature of dense video streams explicitly.

In Turkbeyler and Harris (2010) aerial mosaics are build and geo-located in order to indicate movements on the ground. Observations for the corresponding adjustment task are the image-to-image homographies of a video stream captured by a downward looking airborne camera. The homographies are fixed by Euclidean normalization, i.e. by setting $H_{33} = 1$. The stochastic model assumes additive noise for the algebraic transformation parameters. An adjustment for loop closing is performed with the parameters of the image-to-mosaic transformations for each image as unknowns.[1] For problems of moderate size this leads already to huge band diagonal normal equation matrices with sparse off-diagonal terms. This makes the task of mosaicking several thousands of images computationally intensive, cf. Unnikrishnan and Kelly (2002b).

The detection of loops in an image sequence and another scheme for loop closing is discussed in Caballero et al. (2007): Loops are being hypothesized and detected by considering the pairwise Mahalanobis distances of the estimated positions of the image centers on the mosaic in conjunction with their uncertainties. For the subsequent update of the loops' homographies an optimization procedure is launched by applying and extended Kalman filter. For the subtraction within the filter's update equations a normalization of the homographies is necessary within each update step to fix the scale of the homogeneous representation.

1.3 Contribution

We propose an efficient feature-based method to build large and globally consistent mosaics from video streams captured for instance by a downward looking airborne camera. The approach assumes planar scenes or a fixed projection center. Furthermore, we assume uncalibrated cameras but straight-line preserving optics, i.e. no presence of lens distortion. The approach allows multiple simultaneous loop closing in a batch process after a loop detection stage. Considering the image-to-image homographies as observations, we chose an adjustment model with constraints for theses homographies only, leading to small equation systems to be solved. The involved stochastic model rigorously incorporates the uncertainties of feature extraction, tracking, and matching respectively. By taking advantage of the utilized exponential representation of the homographies the computations become simple and efficient which paves the way for on-board computations.

[1] The authors use the term "bundle adjustment" while our notion on this concept implies the simultaneous estimation of motion parameters (homographies) and structure parameters (image feature positions).

2 Theoretical Background and Modeling

After defining a homography in general we introduce concepts and technical terms for certain homographies needed for the proposed loop closing technique. The discussed adjustment procedure including its model takes advantage of the used minimal parameterization and the resulting simplicity of the constraint equations for the loops to be closed.

2.1 Notation and Preliminaries

Homogeneous vectors are denoted with upright boldface letters, e.g. $\mathbf{x}$ or $\mathbf{H}$, Euclidean vectors and matrices with slanted boldface letters, e.g. $\boldsymbol{l}$ or $\boldsymbol{R}$. For homogeneous coordinates "=" means an assignment or an equivalence up to a common scale factor $\lambda \neq 0$.

For the minimal parameterization of a homography we exploit the power series

$$\exp(\boldsymbol{K}) = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{K}^k = \boldsymbol{I}_3 + \boldsymbol{K} + \frac{1}{2!}\boldsymbol{K}^2 + \dots \tag{1}$$

being the matrix exponential for square matrices analogous to the ordinary exponential function.

For the analytical computations of Jacobians we will use the rule

$$\mathrm{vec}(\boldsymbol{ABC}) = (\boldsymbol{C}^\mathsf{T} \otimes \boldsymbol{A})\mathrm{vec}(\boldsymbol{B}) \tag{2}$$

frequently where the vec operator stacks all columns of a matrix and $\otimes$ denotes the Kronecker product.

2.2 Homographies

Definition. A planar projective transformation is a linear transformation on homogeneous 3-vectors represented by a non-singular 3×3 matrix $\mathbf{H} = (H_{ij})$, cf. Hartley and Zisserman (2000):

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} \tag{3}$$

more briefly, $\mathbf{x}' = \mathbf{H}\mathbf{x}$. The transformation is unique up to scale and has therefore eight degrees of freedom. It can be written in inhomogeneous form as

$$x' = \frac{H_{11}x + H_{12}y + H_{13}}{H_{31}x + H_{32}y + H_{33}} \qquad \text{and} \qquad y' = \frac{H_{21}x + H_{22}y + H_{23}}{H_{31}x + H_{32}y + H_{33}} \tag{4}$$

with $\mathbf{x} = [x, y, 1]^\mathsf{T}$ and $\mathbf{x}' = [x', y', 1]^\mathsf{T}$.

Homographies form a group. Thus, one can "undo" a transformation by computing and applying the inverse transformation (matrix inversion). The concatenation or chaining of two or more transformations results from direct matrix

multiplication. In the following we will use products of homography matrices extensively and will denoted them by an overbar representing sequential chaining of temporal adjacent images within the video stream. The homography matrix for the transformation from image i to image k reads

$$ {}^{k}\overline{\mathbf{H}}_{i} = \prod_{j=i}^{k-1} {}^{j+1}\mathbf{H}_{j} \tag{5} $$

where the product symbol induces matrix multiplications from the left.

Sequential Links and Cross Links. During a flight essentially two kinds of links can occur. For the characterization of these links we adopt the terminology used in Turkbeyler and Harris (2010): Pairwise homographies of consecutive, temporal adjacent images are called *sequential links*. Other homographies are given by image pairs with overlap at crossings after completing a circuit. They constitute the *cross links* of non-temporally adjacent images used for loop closing. Figure 1 shows on the left side an example for a flight with numerous sequential links and a few cross links which can be used to adjust the corresponding loops.

The chaining by homographies can be considered as a topological graph which is non-planar in general. Its cycles constitute the loops. Single connected components are part of no loop. Thus, they will not be affected by the adjustment. The set of cycles that may be used must form an independent and complete cover. Figure 2 shows such a fundamental cycle basis. Of course, choosing the smallest-size cycle basis would reduce the computational costs, cf. Unnikrishnan and Kelly (2002b). But for the sake of simplicity we establish loops along the camera path only.

Cumulative Homographies. In analogy to cumulative sums and cumulative products we define a cumulative homography as the concatenation of homographies starting with the very first sequential homography ${}^{2}\mathbf{H}_{1}$ between image 1

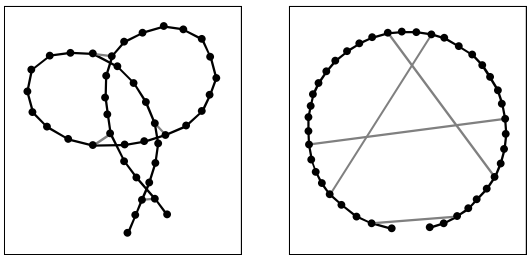

Fig. 1. Left: A hypothetical flight (sequential links in black) with four established cross links (gray) schematically. Right: The corresponding topological graph

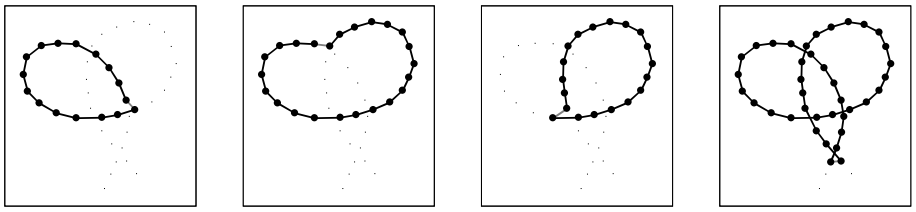

Fig. 2. The cycle basis for the graph depicted in Figure 1 resulting from four detected and established loops along the flight path

and 2:

$$ {}^{i}\overline{\mathbf{H}}_1 = \prod_{k=1}^{i-1} {}^{k+1}\mathbf{H}_k \tag{6} $$

In the following we will extensively use these entities. The subsequence (5) from image i to image k for instance can easily represented by

$$ {}^{k}\overline{\mathbf{H}}_i = {}^{k}\overline{\mathbf{H}}_1 {}^{1}\overline{\mathbf{H}}_i = {}^{k}\overline{\mathbf{H}}_1 \left({}^{i}\overline{\mathbf{H}}_1\right)^{-1} \tag{7} $$

using cumulative homographies only.

2.3 Parameterization

Fixing the scale of a homography in homogeneous representation can be accomplished in various ways: Popular approaches are the fixing of one matrix element or the matrix' Frobenius norm. Fixing one of the elements of $\mathbf{H} = (H_{mn})$, e.g. $H_{33} = 1$, is too restrictive since the case of zero or close to zero entries cannot be excluded in general, cf. (Hartley and Zisserman, 2000, p. 41). Fixing the scale by the Frobenius norm, i.e. $||\mathbf{H}|| = 1$, is frequently used to obtain closed form solutions as this constraint is quadratic in the elements of the matrix $\mathbf{H}$. Both approaches suffer from numerical problems when chaining many transformations as needed for cumulative homographies. Thus, usually a computational expensive re-normalization is necessary after each multiplication. More favorable is the constraint $\det(\mathbf{H}) = 1$ which opens up the vista of minimal representations, too.

Exponential Representation. Minimal representations have recently attract increased attention, cf. Förstner (2010). By avoidance of redundancy additional parameters constraints become superfluous and the resulting equation systems smaller.

A homography matrix can be decomposed in an approximate transformation $\mathbf{H}_0$ and a small, unknown correcting homography $\Delta\mathbf{H}$

$$ \mathbf{H} = \Delta\mathbf{H} \cdot \mathbf{H}_0. \tag{8} $$

This multiplicative expansion facilitates linearization. If both matrices on the r.h.s. have determinant one, the resulting homography matrix has determinant one, too. This can be achieved by using the exponential representation $\Delta\mathbf{H} = \exp(\boldsymbol{K})$, cf. (1), for the homography update. For square matrices $\boldsymbol{A}$ the relation $\det(\exp(\boldsymbol{A})) = \exp(\mathrm{tr}(\boldsymbol{A}))$ holds. Thus, requiring $\det(\Delta\mathbf{H}) = 1$ is equivalent to $\mathrm{tr}(\boldsymbol{K}) = 0$. The matrix

$$\boldsymbol{K} = \begin{bmatrix} k_1 & k_4 & k_7 \\ k_2 & k_5 & k_8 \\ k_3 & k_6 & -k_1 - k_5 \end{bmatrix} \tag{9}$$

is trace-less and depends linearly on the eight parameters $\boldsymbol{k} = [k_1, \ldots, k_8]^\mathsf{T}$ constituting the correction parameters, cf. Begelfor and Werman (2005).

Error Propagation. In the following we will use the exponential representation only. Estimates $(\widehat{\boldsymbol{k}}, \widehat{\boldsymbol{\Sigma}}_{\widehat{k}\widehat{k}}, \mathbf{H}_0)$ for the unknown homography parameters and their corresponding covariance matrix are obtained by given sets of point correspondencies with $\widehat{\boldsymbol{k}} = \mathbf{0}$ since $\mathbf{H}_0$ is updated during the estimation process. The estimated covariance matrix for the nine corresponding homography parameters $\widehat{\boldsymbol{h}} = \mathrm{vec}(\widehat{\boldsymbol{H}})$ is obtained by error propagation for the update transformation (8) with

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{I}_8 \\ -1, 0, 0, 0, -1, 0, 0, 0 \end{bmatrix} \tag{10}$$

since $\mathrm{vec}(\boldsymbol{K}) = \boldsymbol{G}\boldsymbol{k}$ holds for (9). The Jacobian for the error propagation $\widehat{\boldsymbol{\Sigma}}_{\widehat{h}\widehat{h}} = \boldsymbol{J}\widehat{\boldsymbol{\Sigma}}_{\widehat{k}\widehat{k}}\boldsymbol{J}^\mathsf{T}$ is then simply $\boldsymbol{J} = \left(\mathbf{H}_0^\mathsf{T} \otimes \boldsymbol{I}_3\right)\boldsymbol{G}$ being a specialization of the rule (2). For synthetic data we validate this stochastic model in subsection 4.1.

3 Realization

This contribution focuses on the loop closing procedure explicated below in the subsections 3.2 and 3.3. Nevertheless, in the following subsection we present at least an idea how the search area for potential loop closure events can be narrowed down.

3.1 Loop Detection

For the generation of hypothetical loops we consider the Mahalanobis distances between the (error free) image center $\mathbf{x}_c$ and its transformation ${}^j\mathbf{x}_c = {}^j\mathbf{H}_i\mathbf{x}_c$ for each image pair (i, j). The homographies can be obtained from the cumulative ones by (7). Applying the law of error propagation for the relation ${}^j\overline{\mathbf{H}}_1 = {}^j\overline{\mathbf{H}}_i\,{}^i\overline{\mathbf{H}}_1$ we get the uncertainty of the homography ${}^j\overline{\mathbf{H}}_i$. In terms of the covariance matrices ${}^i\overline{\boldsymbol{\Sigma}}_1$ and ${}^j\overline{\boldsymbol{\Sigma}}_1$ of the cumulative homographies it is

$${}^j\overline{\boldsymbol{\Sigma}}_i = \boldsymbol{A}^{-1}({}^j\overline{\boldsymbol{\Sigma}}_1 - \boldsymbol{B}\,{}^i\overline{\boldsymbol{\Sigma}}_1\boldsymbol{B}^\mathsf{T})\boldsymbol{A}^{-\mathsf{T}}, \tag{11}$$

with $\boldsymbol{A} = {}^{i}\overline{\mathbf{H}}_1 \otimes \boldsymbol{I}_3$, $\boldsymbol{B} = \boldsymbol{I}_3 \otimes {}^{j}\overline{\mathbf{H}}_i$, and the corresponding independence assumption.

Hence, the covariance matrix of the predicted image center ${}^{j}\mathbf{x}_c$ in homogeneous coordinates is $\boldsymbol{\Sigma}_{dd} = \boldsymbol{C}\,{}^{j}\overline{\boldsymbol{\Sigma}}_i\boldsymbol{C}^\mathsf{T}$ with $\boldsymbol{C} = \mathbf{x}_c^\mathsf{T} \otimes \boldsymbol{I}_3$ and the Mahalanobis distances read

$$d_{ij}^2 = \mathbf{d}^\mathsf{T}\boldsymbol{\Sigma}_{dd}^{+}\mathbf{d} \qquad \text{with} \quad \mathbf{d} = {}^{j}\mathbf{x}_c - \mathbf{x}_c \tag{12}$$

with spherical normalized entities.

Figure 5 visualizes the discretized reciprocal values of the computed Mahalanobis distances (12) for the example prepared in the introduction and carried out later in subsection 4.2. The image clearly reveals the four loop closing events sketched in Figure 1 by regions of local minima.

3.2 Adjustment Model and Adjustment Procedure

In practice, the number of video frames is large which prohibits approaches estimating all image-to-mosaic homographies. Therefore, we carry out an adjustment for the image-to-image homographies which constitute loops, only. For closing the loops' gaps we chose the adjustment model with constraints between observations only, cf. Koch (1999); McGlone et al. (2004). In the following we consider the estimated correction parameters $\widehat{\boldsymbol{k}} = \mathbf{0}$ for the homographies as observations accompanied by their estimated covariance matrices $\widehat{\boldsymbol{\Sigma}}_{\widehat{k}\widehat{k}}$. Since the number of loops is usually small we get small equation systems, too.

Adjustment Model. Among N observations $\boldsymbol{l}$ we have the G constraints $\boldsymbol{g}(\widehat{\boldsymbol{l}}) = \mathbf{0}$ which have to hold for the true values as well as for the estimated values $\widehat{\boldsymbol{l}}$, namely the fitted observations $\widehat{\boldsymbol{l}} = \boldsymbol{l} + \widehat{\boldsymbol{v}}$ with the estimated corrections $\widehat{\boldsymbol{v}}$. An initial covariance matrix $\boldsymbol{\Sigma}_{ll}^{(0)}$ of the observations is assumed to be known and related to the true covariance matrix $\boldsymbol{\Sigma}_{ll}$ by $\boldsymbol{\Sigma}_{ll} = \sigma_0^2\boldsymbol{\Sigma}_{ll}^{(0)}$ with the possibly unknown scale factor σ_0^2 which can be estimated from the residuals.

Linearization of the constraints by Taylor expansion yields

$$\boldsymbol{g}(\widehat{\boldsymbol{l}}) = \boldsymbol{g}(\boldsymbol{l}_0) + \boldsymbol{J}\widehat{\boldsymbol{\Delta l}} + \ldots \tag{13}$$

with approximate values $\boldsymbol{l}_0$ and $\widehat{\boldsymbol{l}} = \boldsymbol{l} + \widehat{\boldsymbol{v}} = \boldsymbol{l}_0 + \widehat{\boldsymbol{\Delta l}}$. Thus the linear model is $\boldsymbol{g}_0 + \boldsymbol{J}\widehat{\boldsymbol{v}} = 0$ with the contradiction vector $\boldsymbol{g}_0 = \boldsymbol{g}(\boldsymbol{l}_0) + \boldsymbol{J}(\boldsymbol{l} - \boldsymbol{l}_0)$.

The Lagrangian incorporating the method of least squares reads

$$L = \frac{1}{2}\widehat{\boldsymbol{v}}^\mathsf{T}\boldsymbol{\Sigma}_{ll}^{-1}\widehat{\boldsymbol{v}} + \boldsymbol{\lambda}^\mathsf{T}(\boldsymbol{g}_0 + \boldsymbol{J}\widehat{\boldsymbol{v}}) \tag{14}$$

and its minimization w.r.t. the residuals and the Lagrangian multipliers $\boldsymbol{\lambda}$ yields the estimates

$$\widehat{\boldsymbol{v}} = -\boldsymbol{\Sigma}_{ll}\boldsymbol{J}^\mathsf{T}\left(\boldsymbol{J}\boldsymbol{\Sigma}_{ll}\boldsymbol{J}^\mathsf{T}\right)^{-1}\boldsymbol{g}_0 \tag{15}$$

for the estimated corrections being the negative residuals.

Adjustment Procedure. The covariance matrix for observations $\boldsymbol{l} = [\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots]^\mathsf{T}$ is block-diagonal

$$\boldsymbol{\Sigma}_{ll} = \mathrm{Diag}\left(\boldsymbol{\Sigma}_{k_1k_1}, \boldsymbol{\Sigma}_{k_2k_2}, \ldots\right) \tag{16}$$

with 8×8 blocks and has full rank. The size of the matrix to be inverted for the estimation (15) is small for problems of moderate size and complexity. Its size is $8L \times 8L$ with the number of loops L.

During the iterative estimation procedure we do not update the adjusted observations but the approximate homography matrices $\mathbf{H}_0$ by applying $\mathbf{H}_0^{(\nu+1)} = \exp(\widehat{\boldsymbol{K}})\mathbf{H}_0^{(\nu)}$ where ν is the iteration number. The adjusted "observations" $\boldsymbol{k}$ remain zeros.

3.3 Loop Constraints and Jacobians

Loop Constraints. With the observed cross links ${}^k\mathbf{H}_i$ and the observed cumulative homographies ${}^k\overline{\mathbf{H}}_i$ the constraints $\boldsymbol{g}(\boldsymbol{l}_0)$ for each loop are simply

$${}^k\overline{\mathbf{H}}_i - {}^k\mathbf{H}_i = \boldsymbol{O}_{3\times3} \tag{17}$$

with $\det({}^\bullet\mathbf{H}_\bullet) = 1$ for all homography matrices. One has to select eight independent equations. An alternative formulation of the loop constraints is ${}^k\overline{\mathbf{H}}_i\,{}^i\mathbf{H}_k = \boldsymbol{I}_3$ cf. Unnikrishnan and Kelly (2002a), whereby the Jacobian w.r.t. the cross links depends on the sequential links, too.

Jacobians. For the cross links the exponential representation is ${}^k\mathbf{H}_i = \Delta\mathbf{H} \cdot \mathbf{H}_0 \approx \Delta\mathbf{H}(\boldsymbol{I}_3 + \boldsymbol{K})\mathbf{H}_0$ and therefore the Jacobian w.r.t. the parameters $\boldsymbol{k}_i$ is

$$\boldsymbol{J}_c^{(i)} = -(\mathbf{H}_0^\mathsf{T} \otimes \boldsymbol{I}_3)\boldsymbol{G} \tag{18}$$

with (10) and the rule (2). For sequential links we consider the update ${}^k\overline{\mathbf{H}}_i = {}^k\overline{\mathbf{H}}_j \cdot \Delta\mathbf{H}_j\,{}^j\overline{\mathbf{H}}_i$ for a homography as part of a sequence from frame i to k. Expressed in terms of cumulative homographies this reads

$${}^k\overline{\mathbf{H}}_i = {}^k\overline{\mathbf{H}}_1\,{}^1\overline{\mathbf{H}}_j \cdot \Delta\mathbf{H}_j \cdot {}^j\overline{\mathbf{H}}_1\,{}^1\overline{\mathbf{H}}_i = {}^k\overline{\mathbf{H}}_1\left({}^j\overline{\mathbf{H}}_1\right)^{-1} \cdot \Delta\mathbf{H}_j \cdot {}^j\overline{\mathbf{H}}_1\left({}^i\overline{\mathbf{H}}_1\right)^{-1} \tag{19}$$

with the cumulative homographies for the start frames ${}^i\overline{\mathbf{H}}_1$ and the end frames ${}^k\overline{\mathbf{H}}_1$ of each loop. The Jacobians of the cross links w.r.t. the correction parameters $\boldsymbol{k}_j$ are now

$$\boldsymbol{J}_s^{(j)} = \left[\left({}^j\overline{\mathbf{H}}_1\,{}^i\overline{\mathbf{H}}_1^{-1}\right)^\mathsf{T} \otimes \left({}^k\overline{\mathbf{H}}_1\,{}^j\overline{\mathbf{H}}_1^{-1}\right)\right]\boldsymbol{G}. \tag{20}$$

The joint Jacobian is a full matrix for the sequential links augmented by a usually much smaller block-diagonal matrix for the L cross links

$$\boldsymbol{J} = \left[\begin{array}{c|cccc} \boldsymbol{J}_s^{(1)} & \boldsymbol{J}_c^{(1)} & \boldsymbol{O} & \ldots & \boldsymbol{O} \\ \boldsymbol{J}_s^{(2)} & \boldsymbol{O} & \boldsymbol{J}_c^{(2)} & \ldots & \boldsymbol{O} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{J}_s^{(L)} & \boldsymbol{O} & \boldsymbol{O} & \ldots & \boldsymbol{J}_c^{(L)} \end{array}\right]. \tag{21}$$

Its entries depend on cumulative homographies between the start and end image of each loop only.

Note that no special treatment of the homographies not being part of any loop is necessary. The corresponding estimates for the corrections (15) will simply be zeros. No case-by-case analysis is necessary.

4 Experiments

In the following the proposed approach will be empirically evaluated with the help of synthetic data sets and its feasibility will be demonstrated with real data sets.

4.1 Numerical Simulation

To validate the stochastic model of the homography parameterization and to evaluate the performance of the proposed loop closing approach we used simulated data. For each homography we generated a random transformation next to the identity with the eight parameters $\boldsymbol{k} = (k_i)$ being Gaussian distributed according to $k_i \sim N(0, 0.02)$. Then 50 points $\boldsymbol{x}_i$ with coordinates uniformly distributed in $[-1, +1]$ have been transformed according to $\mathbf{x}'_i = \mathbf{H}\mathbf{x}_i$ and $\mathbf{H} = \exp(\boldsymbol{K})$. Gaussian noise with $\sigma = 0.001$ has finally been added to all coordinates.

Figure 3 shows the empirical distribution of the Mahalanobis distance

$$d = \left(\widehat{\boldsymbol{h}} - \widetilde{\boldsymbol{h}}\right)^{\mathsf{T}} \widehat{\boldsymbol{\Sigma}}^{+}_{\widehat{h}\widehat{h}} \left(\widehat{\boldsymbol{h}} - \widetilde{\boldsymbol{h}}\right) \tag{22}$$

for the estimated parameters $\widehat{\boldsymbol{h}} = \mathrm{vec}(\widehat{\mathbf{H}})$ w.r.t. the true homography parameters $\widetilde{\boldsymbol{h}}$ obtained for 1,000 samples. As to be expected, the distribution follows the χ^2_8-distribution.

To specify the performance of the loop closing procedure, we generated various sets of sequential and cross links with loops consisting of 100 sequential links in each case. Figure 4 summarizes the computational times needed for these configurations using non-optimized MATLAB code. The number of iterations has been three for the adjustment procedure. The results show that the computational time needed is proportional to the number of homographies to be considered and to the number of loops to be closed.

4.2 Real Data

For the provision of imagery we abused a virtual globe as camera simulator. This guarantees the model assumptions to be valid; all uncertainties stems from tracking and matching respectively. The camera path corresponds to the topology depicted in Figure 1 with varying height above ground and changing roll, pitch, and yaw angles. During the flight the camera fulfilled a 180-degree-turn

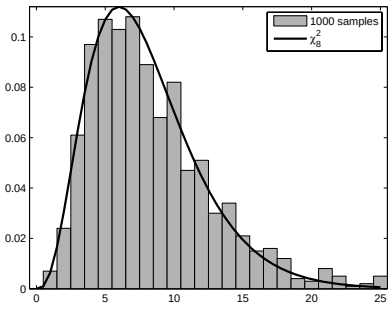

Fig. 3. Empirical distribution of the Mahalanobis distance (22) following the χ^2_8-distribution

		sequential links		
		1,000	2,000	4,000
cross links	2	0.47 sec	0.88 sec	1.71 sec
	4	0.56 sec	1.04 sec	1.97 sec
	8	0.78 sec	1.37 sec	2.50 sec
	16	1.32 sec	2.20 sec	4.00 sec

Fig. 4. Computational times in seconds for various numbers of homographies and loops to be closed. Performed on a standard CPU @ 1.59 GHz with non-optimized MATLAB code.

around its optical axis. Figure 6 shows one of the 1,024 captured images exemplary.

For each image, salient points have been extracted by the Förstner operator (Förstner and Gülch (1987)) and tracked in the corresponding subsequent image by the Lucas-Kanade tracker, cf. Lucas and Kanade (1981). The respective positional uncertainties – represented by covariance matrices – have been determined for incorporation into the adjustment, too. Four loop closing events have been identified by visual inspection. The correspondencies of the cross links have then been established by applying the scale invariant feature transform (Lowe (2004)) in combination with the random sample consensus, cf. Fischler and Bolles (1981).

The mosaic obtained by applying the consecutive homographies only (linear mosaic) is shown in Figure 7 and reveals numerous discrepancies due to the inevitable drift. Figure 7 shows the result of the loop closing procedure proposed in Section 2, too. The computing time was 1.4 seconds on a standard CPU @ 1.59 GHz with non-optimized MATLAB Code in four iterations.

5 Conclusions and Outlook

The proposed loop closing technique offers an easy and efficient way to build large and consistent mosaics quickly in a batch process. It is especially well suited for the processing of video streams since the approach exploits the natural order of these images. The chosen adjustment model, the formulation of the loop constraints, and the exponential parameterization of the homographies lead to efficient computations which can be performed already on board during the flight. The consideration of uncertainties of feature extraction, tracking, and matching makes the approach statistics rigorous.

The capability of instantaneous loop closing compensates for unavoidable drift errors and eases the task of hypotheses testing for potentially occurring

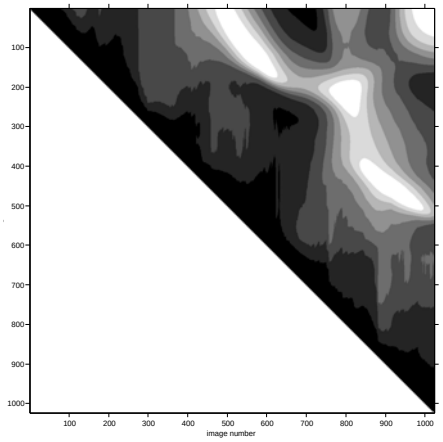

Fig. 5. The visualization of the $I(I-1)/2$ Mahalanobis distances (12) reveals four potential areas of cross links (white spots) for the sequence of I images.

Fig. 6. Exemplary video image of size 320×480 (©GeoContent provided by Google)

loops in the following. Therefore, this technique paves the way for navigation tasks, too, cf. Unnikrishnan and Kelly (2002b) for instance. Another potential application is given in the context of image stabilization or panorama stitching.

For a fully automatic build-up of mosaics an automatic loop detection is necessary. Based on the approach sketched in subsection 3.1 this will be the subject of future investigations. Last but not least the robustness of the approach w.r.t. violated model assumptions — especially non-planarity — should be tested.

Acknowledgments. The author would like to thank Wolfgang Förstner for his inspiration on minimal representations and Eckart Michaelsen for his unweary readiness for discussions.

References

Begelfor, E., Werman, M.: How to put Probabilities on Homographies. IEEE Trans. on Pattern Recognition and Machine Intelligence 27, 1666–1670 (2005)

Caballero, F., Merino, L., Ferruz, J., Ollero, A.: Homography Based Kalman Filter for Mosaic Building. Applications to UAV Position Estimation. In: IEEE International Conf. on Robotics and Automation. pp. 2004–2009 (2007)

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

Förstner, W.: Minimal Representations for Uncertainty and Estimation in Projective Spaces. In: Proc. of Asian Conf. on Computer Vision (2010)

Fig. 7. Left: aerial mosaic built using the estimated 1023 sequential links only. Right: mosaic adjusted by closing the four loops depicted in Figure 2

Förstner, W., Gülch, E.: A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. In: ISPRS Intercommission Workshop, Interlaken (1987)

Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge (2000)

Koch, K.R.: Parameter Estimation and Hypothesis Testing in Linear Models. Springer, Berlin, 2 edn. (1999)

Lowe, D.: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60(2), 91–110 (2004)

Lucas, B.T., Kanade, T.: An Iterative Image Registration Technique with an Application to Stereo Vision. In: Proc. of Image Understanding Workshop. pp. 212–130 (1981)

McGlone, J.C., Mikhail, E.M., Bethel, J. (eds.): Manual of Photogrammetry. American Society of Photogrammetry and Remote Sensing, 5th edn. (2004)

Szeliski, R.: Image alignment and stitching: A tutorial. Foundations and Trends in Computer Graphics and Computer Vision 2(1), 1–104 (2006)

Turkbeyler, E., Harris, C.: Mapping of Movement to Aerial Mosaic with Geo-Location Information. In: Optro 2010 (2010)

Unnikrishnan, R., Kelly, A.: A Constrained Optimization Approach to Globally Consistent Mapping. In: 2002 IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems (IROS '02). vol. 1, pp. 564–569 (2002a)

Unnikrishnan, R., Kelly, A.: Mosaicing Large Cyclic Environments for Visual Navigation in Autonomous Vehicles. In: IEEE International Conf. on Robotics and Automation, 2002 (ICRA '02). vol. 4, pp. 4299–4306 (2002b)