

Probabilistic Analysis of Incremental Light Bundle Adjustment

Vadim Indelman, Richard Roberts and Frank Dellaert

College of Computing, Georgia Institute of Technology, Atlanta, GA 30332, USA

indelman@cc.gatech.edu, richard.roberts@gatech.edu, dellaert@cc.gatech.edu

Abstract

This paper presents a probabilistic analysis of the recently introduced incremental light bundle adjustment method (iLBA) [6]. In iLBA, the observed 3D points are algebraically eliminated, resulting in a cost function with only the camera poses as variables, and an incremental smoothing technique is applied for efficiently processing incoming images. While we have already showed that compared to conventional bundle adjustment (BA), iLBA yields a significant improvement in computational complexity with similar levels of accuracy, the probabilistic properties of iLBA have not been analyzed thus far. In this paper we consider the probability distribution that corresponds to the iLBA cost function, and analyze how well it represents the true density of the camera poses given the image measurements. The latter can be exactly calculated in bundle adjustment (BA) by marginalizing out the 3D points from the joint distribution of camera poses and 3D points. We present a theoretical analysis of the differences in the way that LBA and BA use measurement information. Using indoor and outdoor datasets we show that the first two moments of the iLBA and the true probability distributions are very similar in practice.

1. Introduction

In the past few years, several methods have been proposed for reducing the computational complexity of bundle adjustment (BA). These include methods that exploit the sparsity of the involved matrices in the optimization [13, 10], decoupling the BA problem into several submaps that can be efficiently optimized in parallel [14], constructing a skeletal graph using a small subset of images and incorporating the rest of the images using pose estimation [17], and solving a reduced version of the non-linear system with only part of the camera poses that are carefully chosen so that the reduced system approximates well the full nonlinear problem [11].

Another family of recently suggested methods is

“structure-less” BA [18, 15, 4, 6], in which the camera poses are optimized without including structure parameters into the iterative optimization procedure. In structure-less BA, the 3D points are algebraically eliminated using multi-view constraints and the optimization minimizes the errors in satisfying these constraints, as opposed to optimizing the re-projection errors in conventional BA. If required, all or some of the 3D points can be calculated using standard structure reconstruction techniques based on the optimized camera poses.

The first structure-less BA method was introduced by Steffen et al. [18] who optimized the corrections of image observations subject to satisfying trifocal tensor constraints [2]. A similar concept was developed in [4] using three-view constraints [5] instead of the trifocal tensor. Rodríguez et al. [15] obtained a reduced computational complexity by reformulating the optimized cost function and refraining from correcting the pixels.

Another significant gain in computational complexity was obtained in our previous work [6], that applied a recently developed technique for incremental smoothing [7, 8] to structure-less BA. The developed method, called incremental light bundle adjustment (iLBA), adaptively identifies which camera poses should be optimized each time a new camera is incorporated into the optimization. Typically, incremental optimization involves only a small number of cameras as opposed to always optimizing all cameras in previous structure-less BA methods [18, 15, 4]. iLBA utilizes three-view constraints, which in contrast to using only epipolar constraints in structure-less BA [15], allow consistent motion estimates even when the camera centers are co-linear.

In this paper we present a probabilistic analysis of iLBA. We analyze how well the probability distribution corresponding to the iLBA cost function agrees with the true probability distribution of the camera poses. The latter can be calculated in conventional BA from the joint probability of camera poses and 3D points. An accurate and reliable uncertainties estimate is important in many structure from motion and robotics applications, yet to the best of

our knowledge this is the first time that such an analysis is conducted for structure-less BA methods.

In what follows, we next review the main components in incremental light bundle adjustment. Section 3 then presents a probabilistic analysis of iLBA and Section 4 shows results comparing the iLBA and the true probability distributions of camera poses given image observations. Section 5 concludes and suggests some future research directions.

2. Incremental Light Bundle Adjustment

Incremental light bundle adjustment (iLBA) [6] combines the following two key-ideas: algebraic elimination of 3D points, and incremental smoothing. In this section we review each of these concepts.

2.1. Algebraic Elimination of 3D points

Consider a sequence of N views observing M 3D points, and denote the i^{th} camera pose by x_i and the measured image observation of the j^{th} 3D point l_j by z_i^j . Let also $X \doteq \{x_1^T, \dots, x_N^T\}^T$ and $L \doteq \{l_1^T, \dots, l_M^T\}^T$.

The joint pdf $p(X, L|Z)$ can be explicitly written in terms of the prior information and the actual measurement models:

$$p(X, L|Z) = \text{priors} \cdot \prod_i \prod_j p(z_i^j | x_i, l_j), \quad (1)$$

where $p(z_i^j | x_i, l_j)$ is the measurement model corresponding to the probability density of observing the 3D point l_j from a camera pose x_i at the pixel location z_i^j . Assuming Gaussian distributions, the maximum a posteriori (MAP) estimation

$$X^*, L^* = \arg \max_{X, L} p(X, L|Z),$$

corresponds to the following nonlinear optimization

$$J_{BA}(X, L) = \sum_i \sum_j \left\| z_i^j - \text{proj}(x_i, l_j) \right\|_{\Sigma}^2, \quad (2)$$

where $\text{proj}(\cdot)$ is the projection function [2] for a standard pinhole camera model, and $\|a\|_{\Sigma}^2 \doteq a^T \Sigma^{-1} a$ is the squared Mahalanobis distance with the measurement covariance matrix Σ .

Considering the robot poses that observe some common 3D point l and writing down all the appropriate projection equations, it is possible to algebraically eliminate l , which results in constraints between triplets of poses [19, 6]. One possible formulation of these constraints, recently developed in the context of vision-aided navigation [3, 5], is the

three-view constraints. Assuming three overlapping views k, l and m , these constraints are

$$g_{2v}(x_k, x_l, z_k, z_l) = q_k \cdot (t_{k \rightarrow l} \times q_l) \quad (3)$$

$$g_{2v}(x_l, x_m, z_l, z_m) = q_l \cdot (t_{l \rightarrow m} \times q_m) \quad (4)$$

$$g_{3v}(x_k, x_l, x_m, z_k, z_l, z_m) = (q_l \times q_k) \cdot (q_m \times t_{l \rightarrow m}) - (q_k \times t_{k \rightarrow l}) \cdot (q_m \times q_l) \quad (5)$$

where $q_i \doteq R_i^T K_i^{-1} z$ for any view i and image observation z , K_i is the calibration matrix of this view, R_i represents the rotation matrix from some arbitrary global frame to the i^{th} view's frame, and $t_{i \rightarrow j}$ denotes the translation vector from view i to view j , expressed in the global frame. The first two constraints are the two-view constraints g_{2v} between appropriate pairs of views, while the third constraint, g_{3v} , involves all the three views.

When a 3D point is observed by more than three views, we add a single two-view and three-view constraint between each new view and past views, as further explained in [6].

Consequently, rather than optimizing the bundle adjustment cost function (2), which involves both the pose and landmark variables, in light bundle adjustment (LBA) the cost function is [6]:

$$J_{LBA}(X) \doteq \sum_{i=1}^{N_h} \|h_i(X_i, Z_i)\|_{\Sigma_i}^2, \quad (6)$$

where h_i represents a single two- or three-view constraint ($h_i \in \{g_{2v}, g_{3v}\}$) that is a function of several camera poses $X_i \subset X$ and image observations Z_i in the appropriate views, and N_h is the overall number of such constraints.

One can observe that Eq. (6) indeed does not contain any structure parameters, and hence the overall number of variables in the optimization is significantly reduced compared to the bundle adjustment cost function (2) [6].

2.2. Incremental Smoothing

The second component in iLBA is incremental smoothing [9, 8], which re-calculates only part of the robot's poses each time a new measurement is incorporated into the optimization. Since a detailed exposition of the incremental smoothing approach is beyond the scope of this paper, in this section we only discuss the essentials and refer the reader to [9, 8] for further details.

We represent the probability distributions in BA and LBA formulations using a graphical model, known as the factor graph [12], upon which incremental smoothing is performed. This representation is also used later on in Section 3.

Defining the projection factor for some view x , landmark l and image observation z as

$$f_{\text{proj}}(x, l) \doteq \exp\left(-\frac{1}{2} \|z - \text{proj}(x, l)\|_{\Sigma}^2\right),$$

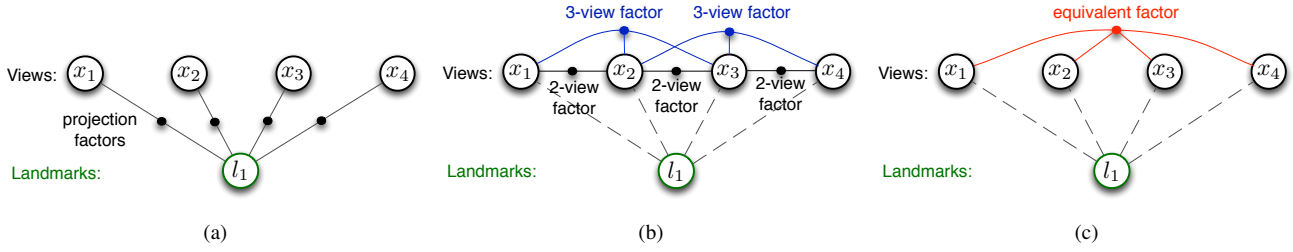


Figure 1: Factor graph formulation for (a) BA and (b) LBA. (c) Factor graph after marginalizing out the landmark l_1 .

and assuming a Gaussian distribution, the joint pdf $p(X, L|Z)$, defined in Eq. (1), can be written as

$$p(X, L|Z) \propto \prod_i \prod_j f_{proj}(x_i, l_j). \quad (7)$$

In a similar manner, the LBA cost function (6) corresponds to some probability distribution $p_{LBA}(X|Z)$, which is formulated next. The following two- and three-view factors are defined:

$$f_{2v}(x_k, x_l) \doteq \exp\left(-\frac{1}{2} \|g_{2v}(x_k, x_l, z_k, z_l)\|_{\Sigma_{2v}}^2\right)$$

and

$$f_{3v}(x_k, x_l, x_m) \doteq \exp\left(-\frac{1}{2} \|g_{3v}(x_k, x_l, x_m, z_k, z_l, z_m)\|_{\Sigma_{3v}}^2\right), \quad (8)$$

where the covariance matrices Σ_{2v} and Σ_{3v} are given in [6].

Taking into account all the available two- and three-view factors, $p_{LBA}(X|Z)$ can be written as

$$p_{LBA}(X|Z) \propto \prod_{i=1}^{N_h} f_{2v/3v}(X_i, Z_i), \quad (9)$$

with $f_{2v/3v} \in \{f_{2v}, f_{3v}\}$ and X_i, Z_i are defined in Section 2. In practice, in order to avoid the trivial solution of zero translation, we normalize each of the constraints g_{2v} and g_{3v} by a translation vector and modify the Jacobian matrices accordingly.

Figures 1a-1b illustrate factor graphs that represent $p(X, L|Z)$ and $p_{LBA}(X|Z)$ in a simple case of 4 camera poses observing 2 different 3D points. Each method uses different factors as discussed above.

In incremental smoothing, the factor graph is converted into a Bayes tree, which is similar to a junction tree but with directed edges, which represents a factorization of the square root information matrix at a given linearization point. Instead of recalculating this factorization from scratch each time a new measurement is added, the factorization from the previous time step is updated by first

adaptively identifying the affected parts in the tree and then recalculating their factorization while re-using calculations elsewhere. Additionally, variables are adaptively re-linearized, such that a MAP estimate up to a tolerance is always available [9, 8].

3. Probabilistic Analysis

This section analyzes how well the LBA distribution $p_{LBA}(X|Z)$ represents the true density $p(X|Z)$. An exact calculation of the latter would marginalize the landmarks from the joint $p(X, L|Z)$

$$p(X|Z) = \int_L p(X, L|Z) dL.$$

While in practice, LBA represents a similar probability density over cameras as BA, there are two root effects that cause the LBA distribution to be an approximation of the true density: First, *LBA discards some mutual information in large camera cliques*, by considering only the mutual information between camera pairs and triplets introduced by them observing the same landmark. Bundle adjustment, on the other hand, induces mutual information between *all* cameras observing the same landmark. Second, *LBA duplicates some information for image measurements used in multiple factors*, double-counting measurements that appear in multiple two- or three-view factors.

As an example of both of these effects, consider observing a landmark l by four views x_1, x_2, x_3 and x_4 , as illustrated in Figure 1. The joint pdf is given by

$$p(X_4, l|Z_4) \propto \prod_{i=1}^4 f_{proj}(x_i, l), \quad (10)$$

where X_4 and Z_4 denote the four camera poses and the four image observations, respectively. On the other hand, the LBA pdf is

$$p_{LBA}(X_4|Z_4) \propto f_{2v}(x_1, x_2) f_{2v}(x_2, x_3) f_{3v}(x_1, x_2, x_3) f_{2v}(x_3, x_4) f_{3v}(x_2, x_3, x_4) \quad (11)$$

which corresponds to the set of two- and three-view factors, as shown in Figure 1.

The first effect, discarding of mutual information, can be seen when comparing the LBA pdf with the pdf resulting from eliminating the landmarks from the BA pdf,

$$p(X_4|Z_4) = \int_{X_4} p(X, L|Z) dX_4 \\ = p(x_1, x_2, x_3, x_4|z_1, z_2, z_3, z_4) \quad (12)$$

The result in the case of BA is a single clique over all cameras. In general, there is no way to exactly factor such a dense clique in a way that reduces complexity. The multiple factors of LBA over pairs and triplets (Eq. (11)) reduce complexity instead by discarding some “links” that would otherwise be introduced between cameras.

The second effect, duplication of some image measurement information, can be seen in the sharing of cameras between LBA factors in Eq. (11). Any two factors sharing a camera in common both use the information from the shared camera, effectively duplicating it. For example, $f_{2v}(x_1, x_2)$ and $f_{2v}(x_2, x_3)$ both use the information from the measurements in camera 2.

As we show in Section 4, despite the above two aspects, the actual LBA distribution is very similar to the true distribution $p(X|Z)$. It is worth mentioning that the presented probabilistic analysis is valid for other existing structure-less BA methods [18, 15, 4] as well.

3.1. Method for Comparing the PDFs of LBA and BA

Because computing the true marginal over cameras for BA $p(X|Z)$ is not tractable in closed form, we use an alternate method to compare the PDFs of LBA and BA. This method evaluates how well LBA and BA agree in both the absolute uncertainty of each camera in a global frame, and the relative uncertainty between all pairs of cameras.

In order to compare uncertainties, we first assume that $p_{LBA}(X|Z)$ and $p(X|Z)$ both are well-approximated as multivariate Gaussian distributions about their MAP estimates

$$p_{LBA}(X|Z) = N(\mu_{LBA}, \Sigma_{LBA}) \\ p(X|Z) = N(\mu, \Sigma).$$

Comparing the accuracy of the MAP estimates themselves was already addressed in [6], where it was demonstrated that LBA yields similar, although a bit degraded, estimates but allow a significant gain in computational complexity. In this paper, we focus on comparing the covariance matrices Σ_{LBA} and Σ . In addition, we present larger datasets that reinforce the previous claim of comparable MAP accuracy between LBA and BA.

In order to compare relative uncertainty between cameras, we compare conditional densities $p(x_i|x_j, Z)$ between all pairs of cameras. This calculation quantifies how well LBA agrees with BA in relative uncertainty, while avoiding calculating the full covariance matrix on all cameras, which quickly becomes intractable for large numbers of cameras. The conditionals are obtained by integrating out all variables other than x_i and x_j ,

$$p(x_i|x_j, Z) = \int_{X \setminus \{x_i, x_j\}, L} p(X, L|Z) / p(x_j|Z).$$

In practice, we do this analytically by approximating the joint as a Gaussian around its MAP estimate, and applying sparse factorization,

$$p(X, L|Z) = p(X \setminus \{x_i, x_j\}, L|x_i, x_j, Z) \\ p(x_i|x_j, Z) p(x_j|Z) \quad (13)$$

from which the desired conditional $p(x_i|x_j, Z)$ can be read off.

4. Results

We use two datasets to evaluate how well the iLBA distribution $p_{LBA}(X|Z)$ represents the true density $p(X|Z)$. In the first dataset (*Cubicle*) the camera observes a cubicle desk in an open space environment from different viewpoints and distances. In the second dataset, *Outdoor*, the camera follows a trajectory encircling a courtyard and building and performing loop closures as shown in Figure 2. Figure 3 shows typical images from these datasets, while Table 1 provides further details regarding the number of views (N) and 3D points (M), as well as the number of total observations in the two datasets.

All methods were implemented using the GTSAM factor graph optimization library¹ [1, 8]. Incremental smoothing was used in all cases, denoted by the prefix *i* (i.e. iLBA and iBA). Image correspondences, as well as the camera calibration matrices, were obtained by first running Bundler² [16] on each dataset. Additional implementation details can be found in [6].

Before discussing probabilistic aspects, we show performance results, in terms of accuracy and computational complexity. As seen in Table 1 and Figure 2, while iLBA yields a similar, but a bit degraded accuracy, the computational cost is significantly lower for iLBA, as compared with BA. A similar observation was made in [6] using smaller datasets.

To eliminate gauge freedoms, the first camera was fixed to the origin and the overall reconstruction scale was determined by setting the range between the first and second

¹<http://tinyurl.com/gtsam>.

²<http://phototour.cs.washington.edu/bundler>.

Dataset	$N, M, \#Obsrv$	iLBA avg. reproj. error	iBA avg. reproj. error
<i>Cubicle</i>	148, 31910, 164358	0.552 pix	0.533 pix
<i>Outdoor</i>	308, 74070, 316696	0.418 pix	0.405 pix

Table 1: Dataset details and performance of iLBA and BA: Re-projection errors and computational cost using incremental smoothing in all methods.

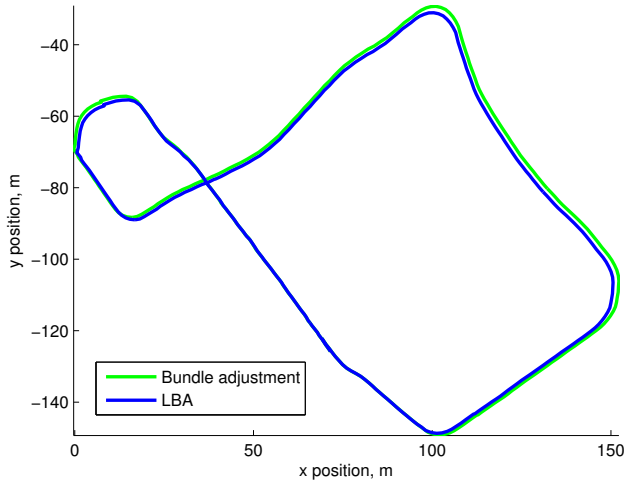


Figure 2: Estimated trajectory in *Outdoor* dataset. LBA and conventional BA produce very similar results.

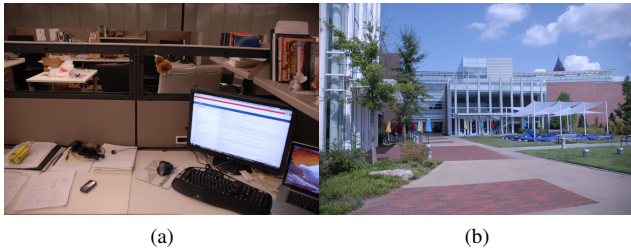


Figure 3: Typical images in the *Cubicle* (a) and *Outdoor* (b) datasets.

camera to some constant number. Consequently, the covariance of the first camera is very small.

We compare the probability density of the cameras estimated by LBA to that of BA by comparing their discrepancy both in the marginal uncertainty of each camera, and in relative uncertainty between each camera pair, as described in Section 3.1. To make these comparisons, we define a discrepancy measure of the square roots of the traces of each

covariance matrix,

$$discrepancy(\Sigma_1, \Sigma_2) \triangleq c \left(\sqrt{\text{tr}(\Sigma_1)} - \sqrt{\text{tr}(\Sigma_2)} \right), \quad (14)$$

where c is a scale factor that converts the unit-less 3D reconstructions into meters, which we determined by physically measuring the dataset collection area, or superimposing the trajectory onto a satellite image. We compute this separately for the blocks of the covariance matrices corresponding to rotation and translation. The units of the discrepancy are radians for rotation ($c = 1$) and meters for translation, with c properly determined to correct the reconstruction scale.

For example, to compare the Gaussian-approximated conditional density of LBA $p_{LBA}(x_i|x_j, Z)$ with covariance $\Sigma_{LBA}^{i|j}$ with that of BA $p(x_i|x_j, Z)$ with covariance $\Sigma_{BA}^{i|j}$, we compute $discrepancy(\Sigma_{LBA}^{i|j}, \Sigma_{BA}^{i|j})$. Similarly for marginals $p_{LBA}(x_i|Z)$ and $p_{BA}(x_i|Z)$, we compute $discrepancy(\Sigma_{LBA}^i, \Sigma_{BA}^i)$. A positive discrepancy value means that the uncertainty estimate of LBA is conservative, whereas a negative discrepancy value means that the uncertainty estimate of LBA is overconfident.

A comparison of the absolute uncertainty for the *Cubicle* dataset is given in Figure 4 and Figures 5a-5b. Figure 4a compares, for each camera pose i , between the covariance trace of Σ_{LBA}^i and Σ_{BA}^i . As seen, the initial uncertainty is very small and it increases as the camera moves around the cubicle deck and drops to low values when the camera captures previously-observed areas thereby providing loop-closure measurements. Figure 4b describes the interaction between the uncertainty of each view and the number of factors that involve this view. As expected, it can be seen that the covariance is higher when less factors are involved and vice versa.

Overall, the absolute uncertainties in LBA and BA are very similar. This can be also observed in Figures 5a-5b that show a histogram of the discrepancy (14) both for position and orientation terms. Typical position discrepancies are near -10^{-4} meters. The discrepancies for relative uncertainties are given in Figures 5c-5d for position and orientation terms.

Figure 6 shows the discrepancy histograms for the *Outdoor* dataset. The absolute and relative discrepancies be-

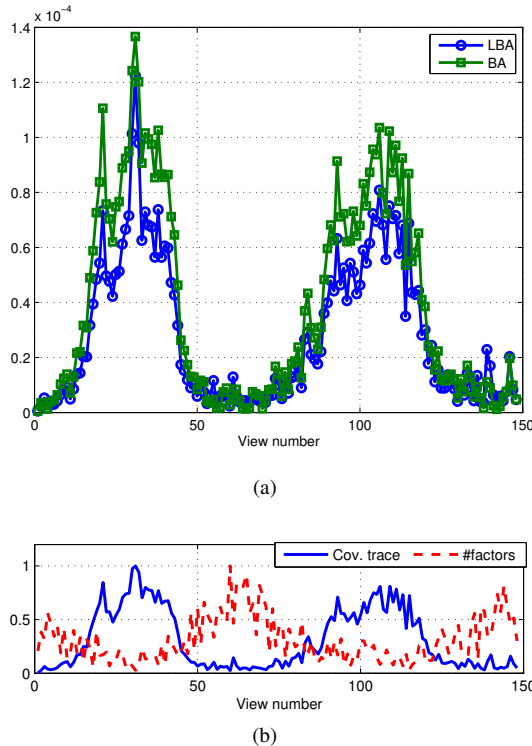


Figure 4: *Cubicle* dataset: (a) Covariance trace of each camera pose. (b) Trace of covariance and number of factors in LBA formulation, both are normalized to 1.

tween LBA and BA are small, e.g. less than 5 centimeters in the absolute position for a trajectory that spans an area of 120×150 meters (cf. Figure 2), and on the order of 10^{-4} radians for the absolute rotation uncertainty.

5. Conclusions

We presented a probabilistic analysis of the recently developed incremental light bundle adjustment (iLBA) method. Two key components of this method are algebraic elimination of the observed 3D points and incremental inference. The first, reduces the number of variables in the optimization, while the second allows to re-calculate only part of the past camera poses when incorporating new measurements. In this paper we analyzed how well the iLBA probability distribution approximates the true distribution of the camera poses given the image observations. The latter can be calculated in bundle adjustment (BA) by marginalizing out the 3D points from the joint density of camera poses and 3D points. The analysis indicated the following two issues that cause the iLBA distribution to be an approximation of the true density: some of the mutual information in large camera cliques is discarded, and some of the image measurements are double-counted. In practice, we showed

using indoor and outdoor datasets, that the effects of these two issues do not lead to significant differences between the iLBA distribution and the true distribution.

References

- [1] F. Dellaert and M. Kaess. Square Root SAM: Simultaneous localization and mapping via square root information smoothing. *Intl. J. of Robotics Research*, 25(12):1181–1203, Dec 2006.
- [2] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [3] V. Indelman. *Navigation Performance Enhancement Using Online Mosaicking*. PhD thesis, Technion - Israel Institute of Technology, 2011.
- [4] V. Indelman. Bundle adjustment without iterative structure estimation and its application to navigation. In *IEEE/ION Position Location and Navigation System (PLANS) Conference*, April 2012.
- [5] V. Indelman, P. Gurfil, E. Rivlin, and H. Rotstein. Real-time vision-aided localization and navigation based on three-view geometry. *IEEE Trans. Aerosp. Electron. Syst.*, 48(3):2239–2259, July 2012.
- [6] V. Indelman, R. Roberts, C. Beall, and F. Dellaert. Incremental light bundle adjustment. In *British Machine Vision Conf. (BMVC)*, September 2012.
- [7] M. Kaess, V. Ila, R. Roberts, and F. Dellaert. The Bayes tree: An algorithmic foundation for probabilistic robot mapping. In *Intl. Workshop on the Algorithmic Foundations of Robotics*, Dec 2010.
- [8] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert. iSAM2: Incremental smoothing and mapping using the Bayes tree. *Intl. J. of Robotics Research*, 31:217–236, Feb 2012.
- [9] M. Kaess, A. Ranganathan, and F. Dellaert. iSAM: Incremental smoothing and mapping. *IEEE Trans. Robotics*, 24(6):1365–1378, Dec 2008.
- [10] K. Konolige. Sparse sparse bundle adjustment. In *BMVC*, September 2010.
- [11] K. Konolige and M. Agrawal. FrameSLAM: from bundle adjustment to realtime visual mapping. *IEEE Trans. Robotics*, 24(5):1066–1077, 2008.
- [12] F. Kschischang, B. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Trans. Inform. Theory*, 47(2), February 2001.
- [13] M. A. Lourakis and A. Argyros. SBA: A Software Package for Generic Sparse Bundle Adjustment. *ACM Trans. Math. Software*, 36(1):1–30, 2009.
- [14] K. Ni, D. Steedly, and F. Dellaert. Out-of-core bundle adjustment for large-scale 3D reconstruction. In *Intl. Conf. on Computer Vision (ICCV)*, Rio de Janeiro, October 2007.
- [15] A. L. Rodríguez, P. E. L. de Teruel, and A. Ruiz. Reduced epipolar cost for accelerated incremental sfm. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 3097–3104, June 2011.
- [16] N. Snavely, S. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3D. In *SIGGRAPH*, pages 835–846, 2006.

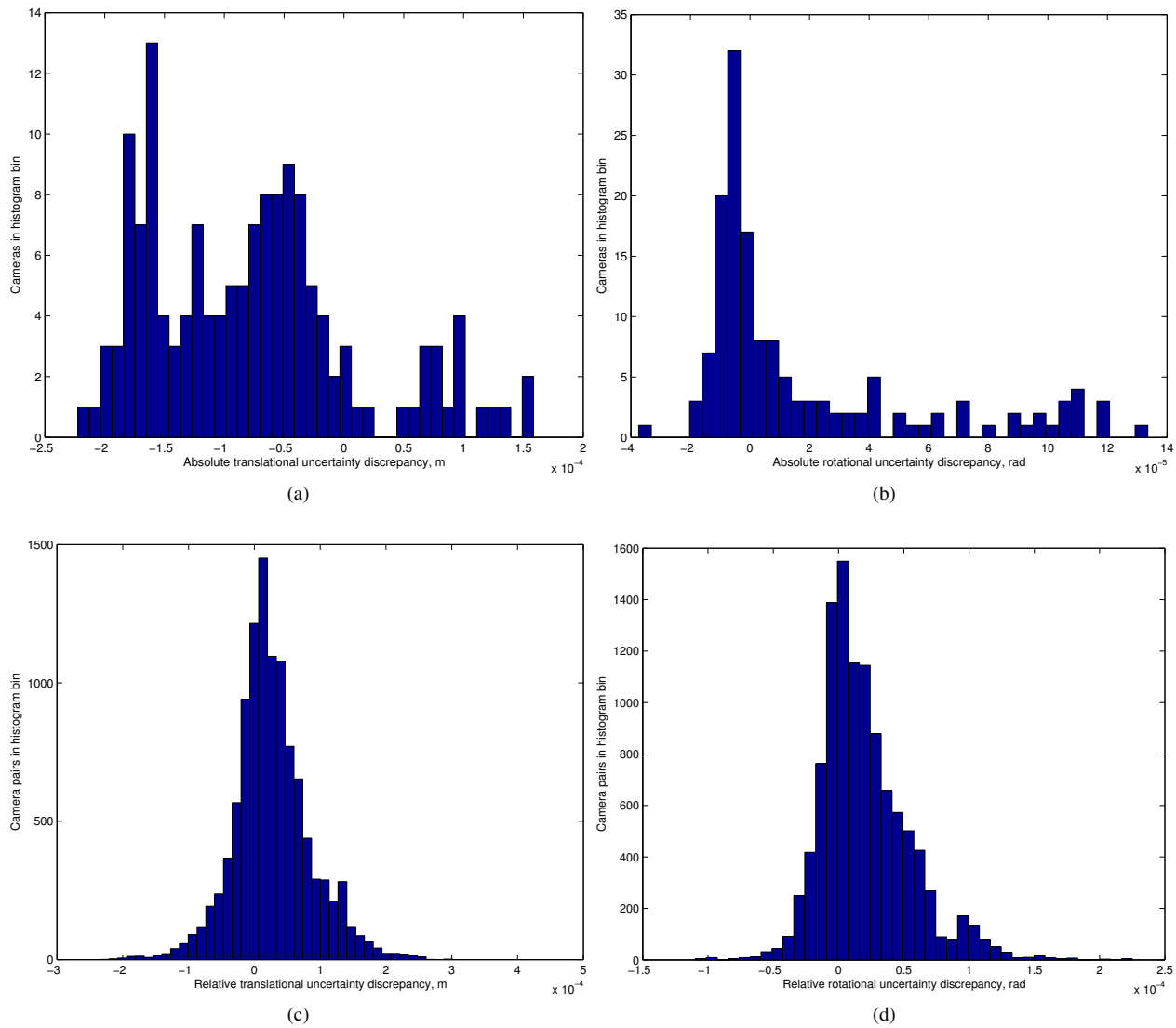


Figure 5: Discrepancy histograms for the *Cubicle* dataset: Absolute position (a) and orientation (b); Relative position (c) and orientation (d) between every camera pair in the sequence.

- [17] N. Snavely, S. M. Seitz, and R. Szeliski. Skeletal graphs for efficient structure from motion. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [18] R. Steffen, J.-M. Frahm, and W. Förstner. Relative bundle adjustment based on trifocal constraints. In *ECCV Workshop on Reconstruction and Modeling of Large-Scale 3D Virtual Environments*, 2010.
- [19] R. Vidal, Y. Ma, S. Soatto, and S. Sastry. Two-View Multi-body Structure from Motion. *Intl. J. of Computer Vision*, 68(1):7–25, 2006.

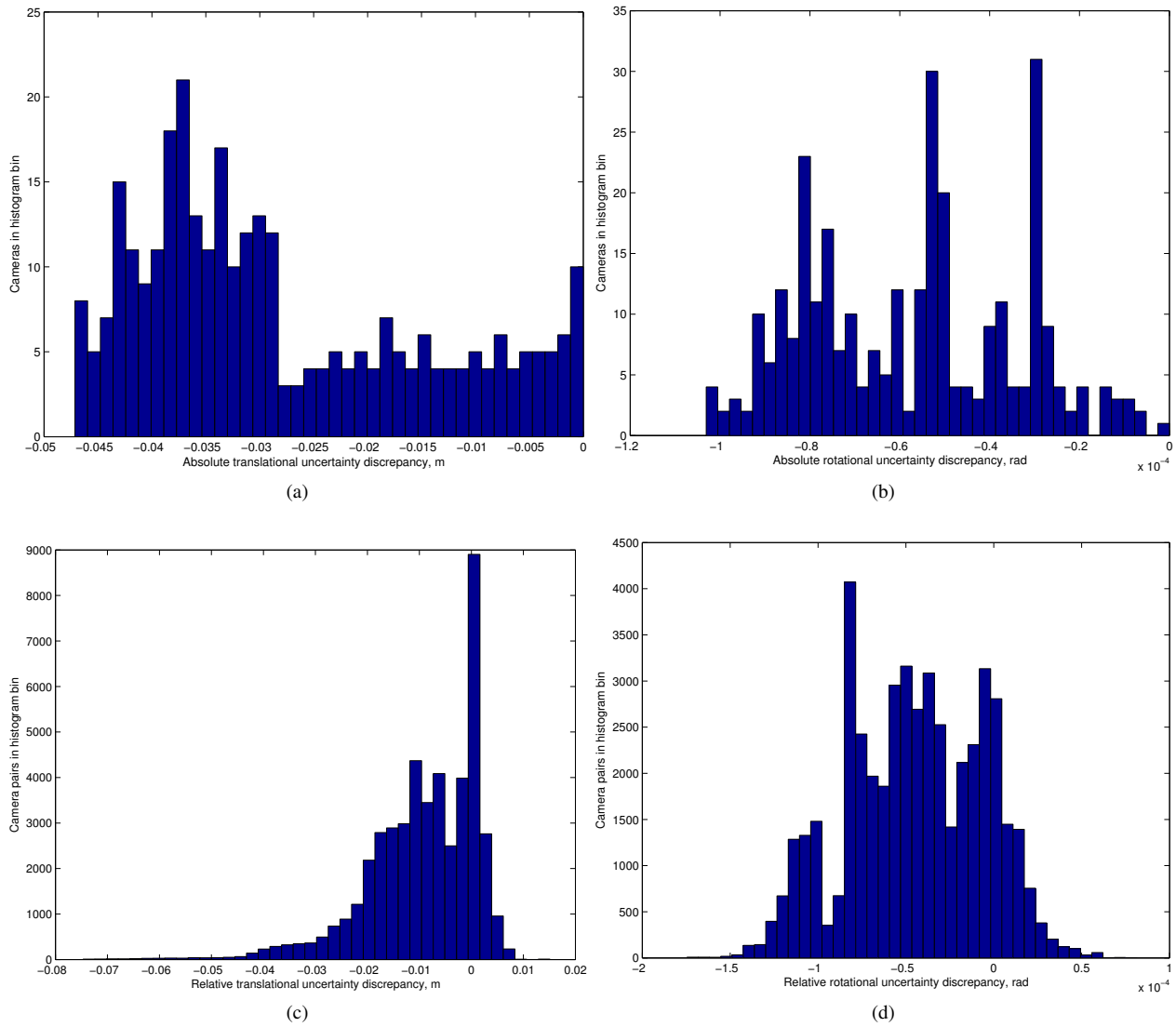


Figure 6: Discrepancy histograms for the *Outdoor* dataset: Absolute position (a) and orientation (b); Relative position (c) and orientation (d) between every camera pair in the sequence.