

Multi-projective Parameter Estimation for Sets of Homogeneous Matrices

Wojciech Chojnacki, Rhys Hill, Anton van den Hengel, Michael J. Brooks

School of Computer Science

The University of Adelaide

Adelaide, SA 5005, Australia

Email: {wojciech.chojnacki, rhys.hill, anton.vandenhengel, michael.brooks}@adelaide.edu.au

Abstract—A number of problems in computer vision require the estimation of a set of matrices, each of which is defined only up to an individual scale factor and represents the parameters of a separate model, under the assumption that the models are intrinsically interconnected. One example of such a set is a family of fundamental matrices sharing an infinite homography. Here an approach is presented to estimating a general set of interdependent matrices defined to within separate scales. The input data is assumed to consist of individually estimated matrices for particular models, which when considered collectively may fail to satisfy the constraints representing the inter-model relationships. Two cost functions are proposed for upgrading, via optimisation, the data of this sort to a collection of matrices satisfying the inter-model constraints. One of these functions incorporates error covariances. Each function is invariant to any change of scale for the input estimates. The proposed approach is applied to the particular problem of estimating a set of fundamental matrices of the form of the example set above. Experimental results are given which demonstrate the effectiveness of the approach.

Keywords—multi-projective parameter estimation, scale independence, maximum likelihood, covariance, homogeneous matrix, fundamental matrix, homography, infinite homography

I. INTRODUCTION

Many of the fundamental tasks in computer vision involve using a matrix that describes some projective geometric relation. It is customary, though somewhat improper, to refer to any such matrix as a homogeneous matrix, meaning that the matrix entries are expressed in homogeneous coordinates.¹ Each non-zero multiple of a homogeneous matrix represents one and the same projective entity, so the matrix is usable and estimable only up to a scale factor. Primary examples of homogeneous matrices are the fundamental matrix and the planar homography matrix, both of size 3×3 . Estimation of these matrices from a set of image-based measurements is a well-researched topic [1].

A number of computer vision problems involve manipulation of not just a single homogeneous matrix, but a whole array of such matrices. Typically, the matrices are intrinsically interconnected. For example, one way to initialise a bundle-adjustment algorithm for projective reconstruction is to use multiple fundamental matrices for which a set of

compatible planar homographies, all induced by a single plane in space, is known [1], [2]. Another example is detection of non-rigid motion based on the use of multiple homographies that are induced by two or more planes in the 3D scene [3]. Multiple homogeneous matrices are multi-projective in nature: they are usable and estimable only to within individual scale factors. In this paper we address the problem of estimating a set of interdependent homogeneous matrices, placing special emphasis on the need to account for the arbitrariness of individual scale factors. A proper treatment of the scale factors, guaranteeing that estimates are not effected by accidental values of these factors, is the main novelty of the present contribution.

Useful sets of multiple homogeneous matrices are subject to intra-set constraints. These are independent from individual-member constraints that may also apply—for example, every member of a set of fundamental matrices satisfies a separate rank-two constraint. For sets of homography matrices different intra-set constraints have been identified in the cases of two and more than two views. Shashua and Avidan [4] have found that homography matrices induced by four or more planes in the 3D scene between two views span a 4-dimensional linear subspace. Chen and Suter [5] derived a set of strengthened constraints for the case of three or more homographies in two views. Zelnik-Manor and Irani [3] have shown that another rank-four constraint applies to a set of so-called relative homographies generated by two planes between four or more views. These latter authors have also derived constraints for larger sets of homographies and views.

In the case of fundamental matrices several intra-set constraints can be identified. One particular set arises in connection with three fundamental matrices linking three views of the same 3D scene. The intra-set constraints in this case consist of three across-the-triple epipolar constraints. In this paper we consider another case where the fundamental matrices between consecutive views share an infinite homography, a situation that occurs when the rotation and calibration of the camera is constant between consecutive views.

Estimation of a set of homogeneous matrices usually proceeds by first generating estimates of individual matrices, and then by upgrading the resulting matrix set to one satisfy-

¹Homogeneous matrices represent, in most cases, non-linear and as such non-homogeneous projective transformations; moreover, their entries typically look rather heterogeneous.

ing the intra-set constraints. Following this pattern, Shashua and Avidan as well as Zelnik-Manor and Irani use low-rank approximation under the Frobenius norm to enforce the rank-four constraint. Chen and Suter enforce their set of constraints also via low-rank approximation, but then employ the Mahalanobis norm with covariances of input homographies. All these estimation procedures involve input matrices coming with specific scale factors. The underlying error measures are such that a change of scale factors may *a priori* result in a different set of estimates.

The aim of this paper is to present an approach for estimating a set of projection transformation matrices which is invariant to any particular choice of individual scale factors. We evolve two cost functions that are scale change insensitive, one of which uses individually estimated matrices as the only available data, and the other additionally employs the covariances of the input matrices. The latter function has a sound statistical basis. We apply both functions to the problem of estimating fundamental matrices that share an infinite homography. While such matrices capture a rather special camera motion, they demonstrate the essence of the approach.

II. MULTI-PROJECTIVE PARAMETER ESTIMATION

Let $\mathbf{X}_1, \dots, \mathbf{X}_I$ be a collection of $k \times l$ matrices. We envisage these matrices as data points in our problem. Each \mathbf{X}_i will be assumed to be known only up to an individual multiplicative non-zero factor. Let $\Theta_1, \dots, \Theta_I$ be another collection of $k \times l$ matrices. We treat these latter matrices as improved versions of the \mathbf{X}_i , still to be determined, that are subject to constraints correlating the entries of the Θ_i across the whole collection. Let $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_I)$ be the composite datum and let $\Theta = (\Theta_1, \dots, \Theta_I)$ be the composite indeterminate. The general problem we are concerned with is to fit Θ to \mathbf{X} so that the intrinsic constraints on Θ are met. Exemplifying this general problem is the following specific problem:

- Fit a set of 3×3 fundamental matrices, with varying epipoles and a common infinite homography, to a given set of 3×3 matrices.

We describe this problem, which will be of particular interest to us, in more detail next.

Suppose that a camera moves in steps producing $I + 1$ views of a 3D scene. For each $1 \leq i \leq I + 1$, let $\mathbf{P}_i = \mathbf{K}_i[\mathbf{R}_i \mid \mathbf{t}_i]$ be the 3×4 projection matrix for the i th view, where the 3×3 rotation matrix \mathbf{R}_i and the length-3 translation vector \mathbf{t}_i represent the Euclidean transformation between the camera and the world coordinate systems, and \mathbf{K}_i is the upper triangular 3×3 matrix which encodes the intrinsic parameters of the camera. Assume that the motion of the camera is special in that the relative orientation of the camera between any two successive views is constant,

$$\mathbf{R}_{i+1}\mathbf{R}_i^{-1} = \mathbf{R} \quad (i = 1, \dots, I),$$

and also that the camera's intrinsic parameters do not vary over the views,

$$\mathbf{K}_i = \mathbf{K} \quad (i = 1, \dots, I + 1).$$

Then, for each $i = 1, \dots, I$, the inter-image homography $\mathbf{H}_{\infty, i, i+1}$ induced by the plane at infinity is constant, given by

$$\mathbf{H}_{\infty, i, i+1} = \mathbf{K}_{i+1}\mathbf{R}_{i+1}\mathbf{R}_i^{-1}\mathbf{K}_i^{-1} = \mathbf{K}\mathbf{R}\mathbf{K}^{-1}.$$

Henceforth we label this common infinite homography as \mathbf{H}_{∞} . For each $i = 1, \dots, I$, the fundamental matrix $\mathbf{F}_{i, i+1}$ relating the i th and $(i + 1)$ th views has the form

$$\mathbf{F}_i = [\mathbf{e}'_i]_{\times} \mathbf{H}_{\infty}.$$

Here \mathbf{e}'_i is the epipole in the $(i + 1)$ th view obtained from the optical centre of the i th view, and, for any length-3 vector \mathbf{a} , $[\mathbf{a}]_{\times}$ is the 3×3 antisymmetric matrix such that $\mathbf{a} \times \mathbf{x} = [\mathbf{a}]_{\times} \mathbf{x}$ for each length-3 vector \mathbf{x} . Let $\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_I)$ be the composite of all the fundamental matrices in question. With $\mathbf{h}_{\infty} = \text{vec}(\mathbf{H}_{\infty})$, where vec denotes vectorisation, $\boldsymbol{\eta} = [\mathbf{e}'_1^T \dots, \mathbf{e}'_I^T, \mathbf{h}_{\infty}^T]^T$, and

$$\boldsymbol{\Pi}(\boldsymbol{\eta}) = (\boldsymbol{\Pi}_1(\boldsymbol{\eta}), \dots, \boldsymbol{\Pi}_I(\boldsymbol{\eta})), \quad \boldsymbol{\Pi}_i(\boldsymbol{\eta}) = [\mathbf{e}'_i]_{\times} \mathbf{H}_{\infty},$$

\mathbf{F} can be represented as

$$\mathbf{F} = \boldsymbol{\Pi}(\boldsymbol{\eta}). \quad (1)$$

Note that \mathbf{F} belongs to at most a $(3I + 9)$ -dimensional subset of the $9I$ -dimensional set of all *a priori* length- I sequences of 3×3 matrices, and that this subset is proper when $I \geq 2$. Thus (1) can be viewed as an intrinsic constraint on \mathbf{F} . Suppose now that an estimate $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_I)$ of \mathbf{F} has been generated in some way. For example, for each i , \mathbf{X}_i might be an estimate of \mathbf{F}_i individually obtained from image data. Our estimation problem is to upgrade \mathbf{X} to $\Theta = (\Theta_1, \dots, \Theta_I)$ so that (1) holds for some $\boldsymbol{\eta}$ and Θ is close to \mathbf{X} in a meaningful sense. The essence of the problem is to find a criterion and effective means for selecting an appropriate $\boldsymbol{\eta}$.

III. COST FUNCTIONS

The general problem of fitting Θ to \mathbf{X} with constraints on Θ is best considered as an optimisation problem. One immediate issue is proper design of a cost function. Since the input matrices are known only up to individual scales, the output matrices should also be determined only to within individual scales. For any I -tuple $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_I)$ of non-zero numbers, set

$$\times \boldsymbol{\lambda} \Theta = (\lambda_1 \Theta_1, \dots, \lambda_I \Theta_I).$$

A correctly defined cost function J should be *multi-homogeneous* in Θ , namely it should obey

$$J(\times \boldsymbol{\lambda} \Theta) = J(\Theta)$$

for each I -tuple λ with non-zero entries. This property guarantees that not just a single Θ but a whole family of equi-valid Θ 's, differing by multi-factors, minimise the cost function. Below we present two relevant multi-homogeneous cost functions, one simple and one more involved.

A. ALS cost function

For each $i = 1, \dots, I$, let

$$\theta_i = \text{vec}(\Theta_i), \quad \mathbf{x}_i = \text{vec}(\mathbf{X}_i),$$

the length of these vectors being kl . The angle α_i between θ_i and \mathbf{x}_i is invariant to multiplying θ_i and \mathbf{x}_i by individual non-zero scalars, and as such gives a natural scale-invariant discrepancy measure between θ_i and \mathbf{x}_i . Taking $\sin \alpha_i$ instead (note that $\sin \alpha_i \approx \alpha_i$ for small α_i) offers a discrepancy measure which is algebraically more manageable. The advantage of the latter choice ensues from the identity

$$\sin^2 \alpha_i = 1 - \cos^2 \alpha_i = 1 - \|\theta_i\|^{-2} \|\mathbf{x}_i\|^{-2} (\theta_i^\top \mathbf{x}_i)^2,$$

which, upon introducing the symmetric projection matrix

$$\mathbf{P}_{\mathbf{x}_i}^\perp = \mathbf{I}_{kl} - \|\mathbf{x}_i\|^{-2} \mathbf{x}_i \mathbf{x}_i^\top,$$

where \mathbf{I}_{kl} is the $kl \times kl$ identity matrix, can be rewritten as

$$\sin^2 \alpha_i = \|\theta_i\|^{-2} \theta_i^\top \mathbf{P}_{\mathbf{x}_i}^\perp \theta_i.$$

Now, a natural cost function for our fitting problem can be put forth as the *algebraic least squares* (ALS) cost function

$$J_{\text{ALS}}(\Theta) = \sum_{i=1}^I \sin^2 \alpha_i = \sum_{i=1}^I \|\theta_i\|^{-2} \theta_i^\top \mathbf{P}_{\mathbf{x}_i}^\perp \theta_i. \quad (2)$$

Note that, for each i , the i th rightmost summand is a quotient of two quadratic forms in θ_i and so is invariant to multiplying θ_i by a non-zero scalar, with the consequence of the overall sum being multi-homogeneous in Θ .

With its inherent virtue as a multi-homogeneous function, the ALS cost function is an *ad hoc* construct lacking a satisfying statistical basis. A statistically sound cost function should at least be capable of properly incorporating *covariance information* regarding the data points \mathbf{X}_i . Equation (2) suggests one immediate modification to obtain a desirably enhanced cost function, and this will be specified next.

B. AML cost function

Continuing to refer to the \mathbf{X}_i 's via their vectorisations, suppose that associated with each \mathbf{x}_i is a $kl \times kl$ raw covariance matrix $\Lambda_{\mathbf{x}_i}^0$. Any such $\Lambda_{\mathbf{x}_i}^0$ is meant to carry the bulk of information about the relative importance of the individual entries of \mathbf{x}_i (see [6], [7] for the raw covariance matrices of homography estimates and [8], [9] and Section VI for the raw covariance matrices of fundamental matrix estimates). Upon upgrading every $\Lambda_{\mathbf{x}_i}^0$ to a corresponding *corrected* covariance matrix

$$\Lambda_{\mathbf{x}_i} = \mathbf{P}_{\mathbf{x}_i}^\perp \Lambda_{\mathbf{x}_i}^0 \mathbf{P}_{\mathbf{x}_i}^\perp,$$

which ensures that the covariance matrix informs about the spread of, specifically, normalised versions of data points around \mathbf{x}_i , one can define an *approximated maximum likelihood* (AML) cost function by setting

$$J_{\text{AML}}(\Theta) = \sum_{i=1}^I \|\theta_i\|^{-2} \theta_i^\top \Lambda_{\mathbf{x}_i}^+ \theta_i,$$

where Λ^+ denotes the Moore–Penrose pseudo-inverse of Λ . As it turns out, the expression for $J_{\text{AML}}(\Theta)$ coincides with the squared *Mahalanobis distance* between any aggregate of normalised, arbitrarily signed variants of the \mathbf{x}_i 's and any aggregate of similar variants of the θ_i 's. Moreover, this distance is an approximation of a more refined, maximum likelihood-based Mahalanobis distance between image-based data points underpinning the generation of the \mathbf{x}_i 's, and respective points on the geometric primitives described by the θ_i 's (cf. [10]–[15]). With the AML cost function justified in this way, when one now takes into consideration the constraints on Θ , the corresponding constrained minimiser of J_{AML} can be regarded as a more accurate, on the average, Θ -estimate than the Θ -estimate embodied by the constrained minimiser of J_{ALS} .

IV. COST FUNCTION OPTIMISATION

Let J be a cost function for fitting Θ to \mathbf{X} of the form

$$J(\Theta) = \sum_{i=1}^I \|\theta_i\|^{-2} \theta_i^\top \mathbf{A}_i \theta_i,$$

where, for each $i = 1, \dots, I$, \mathbf{A}_i is a $kl \times kl$ symmetric matrix. Clearly, the ALS and AML cost functions both conform to that profile. Suppose that the constraints on Θ take the form

$$\Theta = \Pi(\eta), \quad \Pi(\eta) = (\Pi_1(\eta), \dots, \Pi_I(\eta)),$$

where η is a length- d vector. Upon introducing the function

$$J'(\eta) = J(\Pi(\eta)),$$

the constrained optimisation problem in question reduces to that of optimising J' , which is an *unconstrained* optimisation problem.

One way for optimising J' is to use the Levenberg–Marquardt (LM) method. The starting point is to restate J' as

$$J'(\eta) = \sum_{i=1}^I \|\mathbf{f}'_i(\eta)\|^2,$$

where, for each $i = 1, \dots, I$,

$$\begin{aligned} \mathbf{f}'_i(\eta) &= \mathbf{f}_i(\pi_i(\eta)), \\ \mathbf{f}_i(\theta_i) &= \|\theta_i\|^{-1} \mathbf{B}_i \theta_i, \quad \mathbf{B}_i^\top \mathbf{B}_i = \mathbf{A}_i, \\ \pi_i(\eta) &= \text{vec}(\Pi_i(\eta)). \end{aligned}$$

Let $\mathbf{f}'(\boldsymbol{\eta}) = [\mathbf{f}'_1(\boldsymbol{\eta}), \dots, \mathbf{f}'_I(\boldsymbol{\eta})]^\top$. The LM technique makes use of the $Ikl \times (3I + 9)$ Jacobian matrix $\partial_{\boldsymbol{\eta}} \mathbf{f}'$ expressed as

$$\partial_{\boldsymbol{\eta}} \mathbf{f}' = [\partial_{\boldsymbol{\eta}} \mathbf{f}'_1 \mid \dots \mid \partial_{\boldsymbol{\eta}} \mathbf{f}'_I]^\top.$$

For each $i = 1, \dots, I$,

$$\partial_{\boldsymbol{\eta}} \mathbf{f}'_i(\boldsymbol{\eta}) = \partial_{\boldsymbol{\theta}_i} \mathbf{f}_i(\boldsymbol{\pi}_i(\boldsymbol{\eta})) \partial_{\boldsymbol{\eta}} \boldsymbol{\pi}_i(\boldsymbol{\eta})$$

with

$$\partial_{\boldsymbol{\theta}_i} \mathbf{f}_i(\boldsymbol{\theta}_i) = \|\boldsymbol{\theta}_i\|^{-1} \mathbf{B}_i \mathbf{P}_{\boldsymbol{\theta}_i}^\perp, \quad \mathbf{P}_{\boldsymbol{\theta}_i}^\perp = \mathbf{I}_{kl} - \|\boldsymbol{\theta}_i\|^{-2} \boldsymbol{\theta}_i \boldsymbol{\theta}_i^\top.$$

The algorithm iteratively improves on an initial approximation $\boldsymbol{\eta}_0$ to the minimiser of J' by constructing new approximations with the aid of the update rule

$$\boldsymbol{\eta}_{n+1} = \boldsymbol{\eta}_n - [\mathbf{H}(\boldsymbol{\eta}_n) + \lambda_n \mathbf{I}_d]^{-1} \partial_{\boldsymbol{\eta}} \mathbf{f}'(\boldsymbol{\eta}_n)^\top \mathbf{f}'(\boldsymbol{\eta}_n),$$

where $\mathbf{H} = \partial_{\boldsymbol{\eta}} \mathbf{f}'^\top \partial_{\boldsymbol{\eta}} \mathbf{f}'$ and λ_n is a non-negative scalar that dynamically changes from step to step [16].

V. CASE STUDY: ESTIMATING MULTIPLE FUNDAMENTAL MATRICES

We specifically consider the LM-based estimation of multiple fundamental matrices. In this case, for each $i = 1, \dots, I$,

$$\begin{aligned} \boldsymbol{\pi}_i(\boldsymbol{\eta}) &= \text{vec}([\mathbf{e}'_i]_\times \mathbf{H}_\infty) = (\mathbf{H}_\infty^\top \otimes \mathbf{I}_3) \text{vec}([\mathbf{e}'_i]_\times) \\ &= (\mathbf{I}_3 \otimes [\mathbf{e}'_i]_\times) \text{vec}(\mathbf{H}_\infty). \end{aligned}$$

Taking into account that $\text{vec}([\mathbf{e}'_i]_\times) = \mathbf{G} \mathbf{e}'_i$ with \mathbf{G} the 9×3 matrix defined as

$$\mathbf{G} = [\mathbf{k}_6 - \mathbf{k}_8, \mathbf{k}_7 - \mathbf{k}_3, \mathbf{k}_2 - \mathbf{k}_4],$$

where, for each $j = 1, \dots, 9$, \mathbf{k}_j is the length-9 vector whose j th entry is equal to 1 and the remaining entries are all equal to zero, one readily verifies that

$$\begin{aligned} \partial_{\mathbf{e}'_i} \boldsymbol{\pi}_i &= (\mathbf{H}_\infty^\top \otimes \mathbf{I}_3) \mathbf{G}, \quad \partial_{\mathbf{e}'_i} \boldsymbol{\pi}_j = \mathbf{0} \quad (i \neq j), \\ \partial_{\mathbf{h}_\infty} \boldsymbol{\pi}_i &= \mathbf{I}_3 \otimes [\mathbf{e}'_i]_\times. \end{aligned}$$

Representing, for each $i = 1, \dots, I$, $\partial_{\boldsymbol{\eta}} \mathbf{f}'_i$ as

$$\partial_{\boldsymbol{\eta}} \mathbf{f}'_i = [\partial_{\mathbf{e}'_i} \mathbf{f}'_i \mid \dots \mid \partial_{\mathbf{e}'_I} \mathbf{f}'_i \mid \partial_{\mathbf{h}_\infty} \mathbf{f}'_i],$$

one finds furthermore that

$$\begin{aligned} \partial_{\mathbf{e}'_i} \mathbf{f}'_i &= \|\boldsymbol{\pi}_i\|^{-1} \mathbf{B}_i \mathbf{P}_{\boldsymbol{\pi}_i}^\perp (\mathbf{H}_\infty^\top \otimes \mathbf{I}_3) \mathbf{G}, \\ \partial_{\mathbf{e}'_j} \mathbf{f}'_i &= \mathbf{0} \quad (j \neq i), \\ \partial_{\mathbf{h}_\infty} \mathbf{f}'_i &= \|\boldsymbol{\pi}_i\|^{-1} \mathbf{B}_i \mathbf{P}_{\boldsymbol{\pi}_i}^\perp (\mathbf{I}_3 \otimes [\mathbf{e}'_i]_\times). \end{aligned}$$

With $\partial_{\boldsymbol{\eta}} \mathbf{f}'$ thus determined, all that is now needed is a suitable initialisation for the LM method.

The initialisation procedure which we adopt relies upon the solution of the following problem:

- Given $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_I)$ satisfying

$$\mathbf{X} = \times \boldsymbol{\lambda} \mathbf{F} \quad (3)$$

where $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_I)$ has exclusively non-zero entries and $\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_I)$ is such that $\mathbf{F}_i = [\mathbf{e}'_i]_\times \mathbf{H}_\infty$ for each i , express, to within individual scalar factors, all the \mathbf{e}'_i 's and $\mathbf{h}_\infty = \text{vec}(\mathbf{H}_\infty)$ in terms of \mathbf{X} .

A seed $\boldsymbol{\eta}_0 = [\mathbf{e}'_{1,0}, \dots, \mathbf{e}'_{I,0}, \mathbf{h}_{\infty,0}]^\top$ for the LM method is obtained by modifying the solution to the above problem to reflect the fact that the original data set \mathbf{X} admits only an approximate representation along the lines of (3). We omit technical details and merely indicate the steps of the initialisation procedure. These are as follows:

- 1) For each i , let $\mathbf{e}'_{i,0}$ be the left singular vector of \mathbf{X}_i corresponding to the least singular value.
- 2) Select i_0 arbitrarily from the range between 1 and I .
- 3) For each i , let

$$\mathbf{P}_{\mathbf{e}'_{i,0}}^\perp = \mathbf{I}_3 - \|\mathbf{e}'_{i,0}\|^{-2} \mathbf{e}'_{i,0} \mathbf{e}_{i,0}'^\top,$$

for each i, j , let

$$\mathbf{a}_{ij} = \mathbf{X}_i^\top \mathbf{P}_{\mathbf{e}'_{i,0}}^\perp \mathbf{e}'_{j,0},$$

and, for each i , set

$$\lambda_i \lambda_{i_0}^{-1} = -\|\mathbf{a}_{ii_0}\|^{-2} \mathbf{a}_{ii_0}^\top \mathbf{a}_{i_0 i}.$$

- 4) Let

$$\begin{aligned} \mathbf{C} &= [\lambda_1 \lambda_{i_0}^{-1} \mathbf{X}_1^\top \mid \dots \mid \lambda_I \lambda_{i_0}^{-1} \mathbf{X}_I^\top]^\top, \\ \mathbf{E} &= [[\mathbf{e}'_{1,0}]_\times^\top \mid \dots \mid [\mathbf{e}'_{I,0}]_\times^\top]^\top, \end{aligned}$$

and set

$$\mathbf{h}_{\infty,0} = \text{vec}(\mathbf{E}^+ \mathbf{C}).$$

VI. RESULTS

We now present results of experiments aimed at characterising the performance of the proposed estimation methods. The methods were tested on synthetic data to determine their performance under varying amount of noise. Repeated experiments were performed in order to collect results of statistical significance. The regime adopted was to generate true corresponding points for three adjacent stereo configurations and collect performance statistics over many trials in which random Gaussian noise was added to the image points. Many configurations were investigated and the results below are typical. Specifically, we conducted experiments by first choosing a realistic geometric configuration for the camera that executes three moves in space. For each of the three consecutive pairs of views, the relative rotation and calibration of the camera was held fixed, whereas the translation offset was allowed to vary. Next, 100 3D points were randomly selected in the field of view of the camera at each of its four positions, and these were then projected onto four 500×500 pixel images to provide “true” image points. For any of the three consecutive pairs of image planes, each image point in any of the two respective planes was perturbed by independent homogeneous Gaussian noise at a preset level

(for different series of experiments, different noise levels applied). In particular, each true image point shared by two consecutive image pairs was perturbed twice. The resulting three sets of noise-contaminated pairs of corresponding points were ensured to be statistically independent. These three sets were next used as input to FNS, an algorithm for estimating a single fundamental matrix [10], to produce three fundamental matrices \mathbf{X}_i ($i = 1, 2, 3$). Denoting by \mathbf{A}_r^+ the truncated rank- r pseudo-inverse of \mathbf{A} [13], the raw covariance matrix for the vectorisation \mathbf{x}_i of \mathbf{X}_i , based on image points $\{\mathbf{m}_{a,i}, \mathbf{m}'_{a,i}\}_{a=1}^{A_i}$ ² expressed in homogeneous coordinates, was taken to be

$$\Lambda_{\mathbf{x}_i}^0 = (\mathbf{M}_{\mathbf{x}_i})_8^+,$$

where

$$\begin{aligned} \mathbf{M}_{\mathbf{x}_i} &= \|\mathbf{x}_i\|^{-2} \times \\ &\sum_{a=1}^{A_i} (\mathbf{m}_{a,i}^T \mathbf{X}_i^T \Lambda_{\mathbf{m}_{a,i}} \mathbf{X}_i \mathbf{m}'_{a,i} + \mathbf{m}_{a,i}^T \mathbf{X}_i \Lambda_{\mathbf{m}'_{a,i}} \mathbf{X}_i^T \mathbf{m}_{a,i})^{-1} \\ &\quad \times (\mathbf{m}_{a,i} \mathbf{m}_{a,i}^T \otimes \mathbf{m}'_{a,i} \mathbf{m}'_{a,i}^T) \end{aligned}$$

with the image data covariance matrices $\Lambda_{\mathbf{m}_{a,i}}$ and $\Lambda_{\mathbf{m}'_{a,i}}$ assumed to be in their default form $\text{diag}(1, 1, 0)$, corresponding to isotropic homogeneous noise in image point measurement. The matrices \mathbf{X}_i and, independently, the same matrices appended by the covariances $\Lambda_{\mathbf{x}_i}^0$ were next used as input to the ALS and AML methods, respectively, to produce estimates $\hat{\Theta}_{\text{ALS}} = (\hat{\Theta}_{\text{ALS},1}, \hat{\Theta}_{\text{ALS},2}, \hat{\Theta}_{\text{ALS},3})$ and $\hat{\Theta}_{\text{AML}} = (\hat{\Theta}_{\text{AML},1}, \hat{\Theta}_{\text{AML},2}, \hat{\Theta}_{\text{AML},3})$. In addition, a *maximum likelihood* estimate $\hat{\Theta}_{\text{ML}} = (\hat{\Theta}_{\text{ML},1}, \hat{\Theta}_{\text{ML},2}, \hat{\Theta}_{\text{ML},3})$ was generated directly from the image data $\{\mathbf{m}_{a,i}, \mathbf{m}'_{a,i}\}_{i=1}^{A_i}$ by minimising the joint re-projection error

$$\sum_{i=1}^3 \sum_{a=1}^{A_i} (d(\mathbf{m}_{a,i}, \underline{\mathbf{m}}_{a,i})^2 + d(\mathbf{m}'_{a,i}, \underline{\mathbf{m}}'_{a,i})^2)$$

over all points $\mathbf{m}_{a,i} = \mathcal{N}(\mathbf{P}\mathbf{M}_{a,i})$ and $\mathbf{m}'_{a,i} = \mathcal{N}(\mathbf{P}'_i\mathbf{M}_{a,i})$, where $\mathbf{P} = [\mathbf{I}_3 \mid \mathbf{0}]$, $\mathbf{P}'_i = [\mathbf{H}_\infty \mid \mathbf{e}_i]$ for each i , and, $\mathbf{M}_{a,i}$ is a point in 3D space for any a and i , with \mathbf{H}_∞ and the \mathbf{e}_i 's and $\mathbf{M}_{a,i}$'s free to vary. With $\mathbf{m} = [m_1, m_2, m_3]^T$, $\mathcal{N}(\mathbf{m}) = \mathbf{m}/m_3$ is a normalisation procedure whose application ensures that the third coordinate of a given planar point is unity, and $d(\mathbf{m}, \mathbf{n})$ denotes the Euclidean distance between the planar points \mathbf{m} and \mathbf{n} that have been normalised in the above sense. The initial \mathbf{H}_∞ and \mathbf{e}_i 's were obtained by using the procedure described in the previous section. Each initial $\mathbf{M}_{a,i}$ was obtained by triangulating from $\mathbf{m}_{a,i}$ and $\mathbf{m}'_{a,i}$ with \mathbf{P} and \mathbf{P}'_i defined in terms of the initial \mathbf{H}_∞ and \mathbf{e}_i . The \mathbf{e}_i 's, \mathbf{H}_∞ , and $\mathbf{M}_{a,i}$'s were then recomputed iteratively by an LM scheme adapted to the task of minimising the joint re-projection error. With $\hat{\Theta}_{\text{ML},i}$'s,

²In some experiments, to avoid singular configuration, the number A_i of pairs of corresponding points was less than 100.

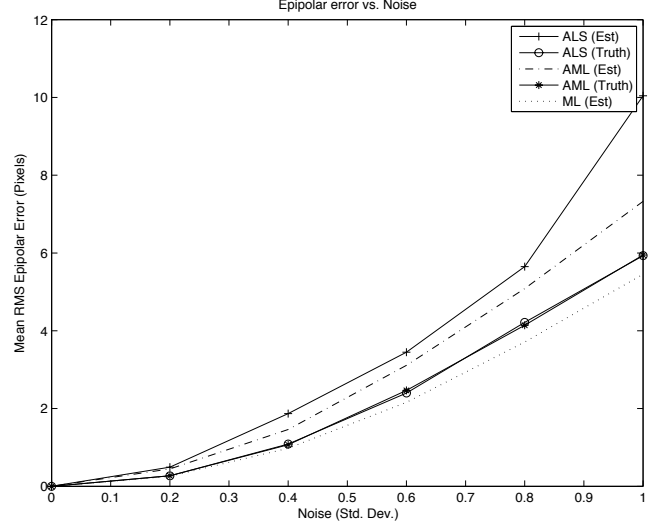


Figure 1. Comparison of three estimation methods.

$\hat{\mathbf{H}}_{\text{ML},\infty}$, and $\hat{\mathbf{M}}_{\text{ML},a,i}$'s denoting the terminal values, the estimates $\hat{\Theta}_{\text{ML},i}$ were finally evolved as

$$\hat{\Theta}_{\text{ML},i} = [\hat{\mathbf{e}}_{\text{ML},i}] \times \hat{\mathbf{H}}_{\text{ML},\infty}.$$

A comparison of the ALS, AML, and ML methods is shown in Figure 1. For $\hat{\Theta} = (\hat{\Theta}_1, \hat{\Theta}_2, \hat{\Theta}_3)$ generated by any of these techniques, the common distance used to quantify data-model discrepancies was the *symmetric epipolar error*

$$\sum_{i=1}^3 \sum_{a=1}^{A_i} (d(\mathbf{m}'_{a,i}, \hat{\Theta} \mathbf{m}_{a,i})^2 + d(\mathbf{m}_{a,i}, \hat{\Theta}^T \mathbf{m}'_{a,i})^2),$$

where d now denotes the Euclidean distance between a point and a line. The ALS and AML schemes were seeded both with true values of \mathbf{H}_∞ and \mathbf{e}_i 's, and with estimates produced by the initialisation procedure described in the previous section. Upon inspection, it is found that the AML scheme produces estimates with very competitive epipolar error values in comparison to the ML method, with the ALS technique lagging further behind, which matches expectations.

VII. CONCLUSION

This work proposes an approach to estimating a set of homogeneous matrices that are subject to individual-member and intra-set constraints. Two cost functions are presented for upgrading individually estimated matrices, which do not necessarily satisfy the underlying constraints, to matrices satisfying the constraints. Each function is invariant to possible changes of individual scales of the input matrices. This scale invariance property is a novel element of design, differentiating the proposed functions from ones previously proposed for similar purposes, not guaranteeing to produce estimates independent of accidental scales of the input. One

of the functions uses the individually estimated matrices as sole input, the other incorporates also the covariances of the input matrices. When optimised, the latter function produces a set of matrix estimates that approximate the maximum likelihood estimates of the matrix set. The proposed approach is tested on the problem of estimating a set of fundamental matrices that share an infinite homography. The Levenberg-Marquardt algorithm evolved to optimise the covariance-enabled cost function for this case produces results showing that the approach is feasible and efficient.

ACKNOWLEDGEMENT

This research was partially supported by the Australian Research Council.

REFERENCES

- [1] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge: Cambridge University Press, 2004.
- [2] R. Kaucic, R. I. Hartley, and N. Y. Dano, "Plane-based projective reconstruction," in *Proc. Eighth Int. Conf. Computer Vision*, vol. 1, 2001, pp. 420–427.
- [3] L. Zelnik-Manor and M. Irani, "Multiview constraints on homographies," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 2, pp. 214–223, 2002.
- [4] A. Shashua and S. Avidan, "The rank 4 constraint in multiple (≥ 3) view geometry," in *Proc. 4th European Conf. Computer Vision*, ser. Lecture Notes in Computer Vision, vol. 1065, 1996, pp. 196–206.
- [5] P. Chen and D. Suter, "Rank constraints for homographies over two views: revisiting the rank four constraint," *Int. J. Computer Vision*, vol. 81, no. 2, pp. 205–225, 2009.
- [6] K. Kanatani, N. Ohta, and Y. Kanazawa, "Optimal homography computation with a reliability measure," *IEICE Trans. Inf. Syst.*, vol. E83-D, no. 7, pp. 1369–1374, 2000.
- [7] P. Chen and D. Suter, "Error analysis in homography estimation by first order approximation tools: a general technique," *J. Math. Imaging Vision*, vol. 33, no. 3, pp. 281–295, 2009.
- [8] P. H. S. Torr and A. Zisserman, "Performance characterization of fundamental matrix estimation under image degradation," *Machine Vision and Applications*, vol. 9, pp. 321–333, 1997.
- [9] K. Kanatani, "High accuracy fundamental matrix computation and its performance evaluation," in *Proc. 17th British Machine Vision Conf.*, vol. 1, 2006, pp. 217–226.
- [10] W. Chojnacki, M. J. Brooks, A. van den Hengel, and D. Gawley, "On the fitting of surfaces to data with covariances," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1294–1303, 2000.
- [11] —, "From FNS to HEIV: A link between two vision parameter estimation methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 264–268, 2004.
- [12] —, "FNS, CFNS and HEIV: A unifying approach," *J. Math. Imaging Vision*, vol. 23, no. 2, pp. 175–183, 2005.
- [13] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*. Amsterdam: Elsevier, 1996.
- [14] Y. Leedan and P. Meer, "Heteroscedastic regression in computer vision: problems with bilinear constraint," *Int. J. Computer Vision*, vol. 37, no. 2, pp. 127–150, 2000.
- [15] B. Matei and P. Meer, "A general method for errors-in-variables problems in computer vision," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, 2000, pp. 18–25.
- [16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*. Cambridge: Cambridge University Press, 1995.