Optimization on the Manifold of Multiple Homographies

Anders Eriksson, Anton van den Hengel
University of Adelaide
Adelaide, Australia
{anders.eriksson, anton.vandenhengel}@adelaide.edu.au

Abstract

It has long been known that the set of homographies for several planes in two images, as well as the homologies of a pair of planes in several images, all lie in a 4-dimensional subspace. It has also been shown that enforcing these constraints improves the accuracy of the homography estimation process.

In this paper we show that the constraints on such collections of homographies are actually stronger than was previously thought. We introduce a new way of characterizing the set of valid collections of homographies as well as suggest a computationally efficient optimization scheme for minimizing over this set. The proposed method, a generalization of Newton's method to manifolds, is experimentally demonstrated on a number of example scenarios with very promising results.

1. Introduction

This paper deals with the problem of determining the geometric relationship between the image motion of planar surfaces across multiple perspective views. In a pair of images this motion, induced by a single plane, is captured by a homography, a 2-dimensional projective transformation. It has long been known that for multiple views and/or multiple planes there exist linear subspace constraints on the induced homographies. It was shown in [14] that a collection of four or more homographies between two views lie in a 4-dimensional subspace. The work in [16] built on these results to show that the relative homographies (homologies) between two planes in multiple views span a 4-dimensional subspace as well. They also extended these findings and proved that similar rank constraints exist in the case of a greater number of planes and views. In this paper we revisit these types of constraints and show that the above rank constraints on the homographies can actually be strengthened.

The task of estimating a single homography given measured images points is well understood, [7, 6, 12, 11]. However, multiple homography estimation has not received equal attention. Given multiple images of a scene containing multiple planar surfaces one obvious approach would be to estimate each of the induced homographies individually. In the absence of noise the collection of such homographies will be compatible with a valid scene configuration and all constraints will be fulfilled. However, if the measurements has been corrupted by noise this will not be the case. The problem we address in this paper is the de-noising of a collection of estimated homographies by ensuring that they agree with a valid scene. Imposing these constraints is not only useful for improving homography estimations but also providing sets of compatible homographies allows for them to be used in applications such as mosaicing, reconstruction and self-calibration.

Shashua and Avidan in [14] showed that any collection of homographies, induced by planes in two views, will lie in a 4-dimensional subspace, and provided a means of estimating this subspace through a truncated SVD projection. However, even though it was shown that this improved the estimated homographies, it is neither the optimal nor correct way in which to include such constraints.

A number of improvements upon this method have been devised, including that of Chen and Suter [4] which derived a set of strengthened constraints similar to ours, but limited to the case in which there are three or more homographies in two views. They also identified some additional and very interesting properties of these constraints, however, this did not lead to a very computationally efficient minimization scheme. Malis and Cipolla [13], proposed the use of rank constraints on collections of collineations between planar surfaces in an image sequence. These constraints are not sufficient, so in order to ensure valid collineations a post processing stage in the form of a least squares fitting is added.

In this work we introduce a novel way of characterizing the constraints on valid collection of homographies as well as a computationally efficient optimization scheme for applying them. We prove that the set of valid collections of homographies forms a smooth manifold $\mathcal{M}$, and that by generalizing the standard Newton's method to a manifold

setting we arrive at a new gradient descent method, actually a damped Newton method, on $\mathcal{M}$. The aim of this paper is not to introduce a state of the art algorithm for estimating homographies but rather to try and contribute to the understanding of the constraints on multiple homographies and how to enforce them efficiently. Our approach is deliberately kept as general as possible, and the proposed algorithm will work on any twice differentiable objective function.

We show that including these strengthened constraints on the homographies not only ensures that they are compatible with a valid scene configuration, but also significantly improves the accuracy of the estimated homographies. In addition, there is no lower limitation on the number of homographies required, as the case is in previous work. We can improve the estimation for as few as two homographies.

We have chosen to illustrate our method on the simplest possible problem formulation. Given a number of individually estimated homographies $\tilde{H}$ we wish to identify the collection of improved estimates H that fulfills all of the implied constraints and is as close as possible to $\tilde{H}$. That is, the minimization problem

$$\min_{H} \quad f(H) = ||\tilde{H} - H|| \tag{1}$$

$$\text{s.t.} \quad H \in \{\text{all valid collections of homographies}\}, \tag{2}$$

where $||\cdot||$ can be any (twice-differentiable) seminorm. The proposed method is demonstrated on a number of synthetic as well as real-world scenarios.

2. Preliminaries

2.1. Homographies

First, we introduce the notation and some basic properties of homographies and homologies.

Let $Q = \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T$ denote the homogeneous 3D coordinates of a scene point and $q = \begin{bmatrix} x & y & 1 \end{bmatrix}^T$ the homogeneous coordinates of an image point. The projection matrix of camera i is written $P_i = \begin{bmatrix} A_i & | & a_i \end{bmatrix}$. We assume that $P_1 = \begin{bmatrix} I & | & 0 \end{bmatrix}$. A planar surface π_j is defined by $\pi_j = \begin{bmatrix} -v_j^T & 1 \end{bmatrix}$, so that a point Q is on π_j if and only if $\pi_j^T Q = 0$.

The homography induced by a single plane π_j from view $P_2 = \begin{bmatrix} A & | & a \end{bmatrix}$ (dropping subscripts here for clarity) to view $P_1 = \begin{bmatrix} I & | & 0 \end{bmatrix}$ is

$$h_j = A + a v_j^T. \tag{3}$$

Now, for multiple planes in two views. We let H denote the collection of m homographies between two views, induced by planes $\pi_1, ..., \pi_m$. That is, the j^{th} column of H corresponds to the vectorization[1] of homography h_j.

$$H = \begin{bmatrix} \vec{h}_1 & ... & \vec{h}^m \end{bmatrix}_{9\times m}. \tag{4}$$

Since $\vec{h}_j$ can be written

$$\vec{h}_j = \text{vec}(h_j) = \text{vec}(A + a v_j^T) = \vec{A} + (I_3 \otimes a) v_j, \tag{5}$$

(where $\otimes$ denotes the Kronecker product) we arrive at

$$H = \begin{bmatrix} \vec{A} + (I_3 \otimes a)v_1 & ... & \vec{A} + (I_3 \otimes a)v_m \end{bmatrix} = \tag{6}$$

$$= \begin{bmatrix} a & 0 & 0 & \\ 0 & a & 0 & \vec{A} \\ 0 & 0 & a & \end{bmatrix}_{9\times 4} \begin{bmatrix} v_1 & ... & v_m \\ 1 & ... & 1 \end{bmatrix}_{4\times m}. \tag{7}$$

As the homographies in $\tilde{H}$ are only known up to a scale, we need to parametrise scaled versions of H. This is obtained by

$$H = \begin{bmatrix} a & 0 & 0 & \\ 0 & a & 0 & \vec{A} \\ 0 & 0 & a & \end{bmatrix}_{9\times 4} \begin{bmatrix} v_1 & ... & v_m \\ \lambda_1 & ... & \lambda_m \end{bmatrix}_{4\times m} = UV. \tag{8}$$

By extending this formulation to more $(N > 2)$ views we see that H^N, the multiple view equivalent of H, can be expressed as

$$H^N = \begin{bmatrix} a_1 & 0 & 0 & \\ 0 & a_1 & 0 & \vec{A}_1 \\ 0 & 0 & a_1 & \\ & \vdots & & \\ a_N & 0 & 0 & \\ 0 & a_N & 0 & \vec{A}_N \\ 0 & 0 & a_N & \end{bmatrix}_{9N\times 4} \begin{bmatrix} v_1 & ... & v_m \\ \lambda_1 & ... & \lambda_m \end{bmatrix}_{4\times m} \tag{9}$$

The observation made in [14] was that as a consequence of equations (8) and (9) the rank of both H and H^N is at most 4. We show in Theorem 1 below that stronger constraints than this may actually be implied on the basis of these equations, but before we doing so we also consider the constraints on sets of relative homographies.

In [16] constraints similar to those above were derived for the case in which two planes appeared in multiple views. The constraints in that case apply to the set of homologies (or relative homographies [7]) induced by the planes.

For a sequence of m images $J_1, ..., J_m$ of two planes π_1 and π_2, if h_j^i represents the homography induced by plane π_i between a reference image J_1 and image J_j, then the relative homography, or planar homology, between the two planes in image J_j is the composition $\bar{h}_j = (h_j^1)^{-1} h_j^2$. If we let $\bar{H}$ denote the homology equivalent of the homography composition H over the m images, it can be shown that

$$\bar{H} = \begin{bmatrix} \beta_1 I & \\ \beta_2 I & I \\ \beta_3 I & \end{bmatrix}_{9\times 4} \begin{bmatrix} \alpha_1 & ... & \alpha_m \\ 1 & ... & 1 \end{bmatrix}_{4\times m} \tag{10}$$

[1] If $M = \left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]$, then $\text{vec}(M) = \vec{M} = [\, a \; b \; c \; d \,]^T$.

where α and β represent plane-dependent and view-dependent components of the factorization of $\bar{H}$. We refer to [16] for the details. The form of this composition and that of H and H^N show that a valid collection of homographies or homologies is not a general rank-4 matrix. The additional constraint provided by equations (8)-(10) also require that the first term in the factorization must lie in some linear subspace, $\mathcal{L}$. This is a key observation, and it is this restriction that provides the improved constraints on collections of homographies. We will later on show that the set of matrices on this form actually make up a low-dimensional manifold embedded in Euclidean space.

For the remainder of the paper we describe our results in terms of multiple homographies in two views only, and thus base our analysis on equation (8). Note, however, that all results also hold for multiview ($N > 2$) homographies (equation (9)), as well as homologies (equation (10)), and that generalizing them is a very straight forward exercise. We will return to the characterization of the set of valid collection of homographies in section 4.1. Next we instead turn to methods for optimizing on manifolds.

3. Optimization on Manifolds

The field of constrained optimization is a well established area of research, with a large number of numerical tools available. This standard approach does, however, require that one can readily describe both objective functions and constraints in a canonical form, something that is not always straight forward.

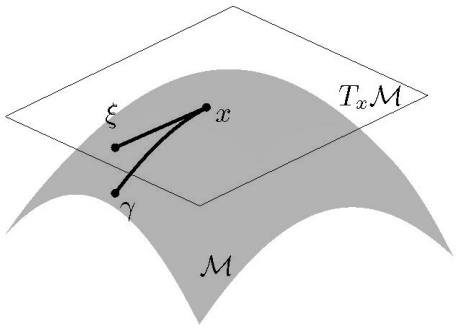


Figure 1. Optimization on a manifold. The tangent space $T_{\mathcal{M}}(x)$, search direction ξ and curve $\gamma(t)$ at a point $x \in \mathcal{M}$.

In this section we give a brief description of a powerful alternative to these classical approaches, namely geometrically constrained optimization or optimization on manifolds, [5, 10, 8]. The underlying idea behind this special class of problems is that constraints can be implicitly enforced by requiring that the solution of the optimization problem lies on a smooth manifold. Such a formulation typically results in a problem of lower dimension and thus hopefully with better numerical properties than an equivalent standard constrained optimization approach. This technique has also been described as unconstrained optimization in a constrained search space.

A manifold can be thought of as a smooth surface in $\mathbb{R}^n$ without folds or tears and that does not self-intersect (see figure 1). Globally, its structure may be highly complex. However, locally it resembles the Euclidean space $\mathbb{R}^d$, where $d \leq n$ is the dimension of the manifold. The tangent space $T_{\mathcal{M}}(x)$ of $\mathcal{M}$, at point x, is a real vector space that contains all possible directions in which one can pass through x on $\mathcal{M}$. It is this local, linear structure that is exploited in order to generalize standard methods for unconstrained optimization to work on manifolds.

Steepest descent methods for minimizing a function $f(x)$ in Euclidean spaces use a search direction ξ at a point x_k obtained by following the negative gradient,

$$\xi = -\nabla f. \tag{11}$$

In Newton's method the gradient is pre-multiplied by the inverse of the Hessian of f at x_k,

$$\xi = -(\text{hess} f)^{-1} \nabla f. \tag{12}$$

The update is then found by carrying out a line search in the direction of ξ, finding the step length t^* that minimizes $f(x_k + t\xi)$, and letting $x_{k+1} = x_k + t^*\xi$.

When minimizing on a non-Euclidean manifold the idea is to use the tangent space, (the local linear coordinate system around a point x) to define a gradient and Hessian at that point on the manifold. And instead of searching along straight lines in a Euclidean space, we then follow curves on $\mathcal{M}$. We have

Definition 1. *The gradient of $f(x)$ on a manifold $\mathcal{M}$ is defined as the unique element $\nabla f \in T_{\mathcal{M}}(x)$ for which*

$$< \nabla f, y >= \frac{d}{dt} f(x + ty)\bigg|_{t=0}, \tag{13}$$

holds for all $y \in T_{\mathcal{M}}(x)$.

Definition 2. *The Hessian of $f(x)$ on $\mathcal{M}$ is defined as the linear operator* $\text{hess} f : T_{\mathcal{M}}(x) \mapsto T_{\mathcal{M}}(x)$ *that satisfies*

$$< \text{hess} f[y], y >= \frac{d^2}{dt^2} f(x + ty)\bigg|_{t=0}, \ \forall y \in T_{\mathcal{M}}(x). \tag{14}$$

Using these definitions we can subsequently obtain a descent direction ξ as in (11) or (12).

What remains is then to define the curves γ on $\mathcal{M}$ to carry out the line searches along. Finding such curves is not always straight forward and can lead to computationally expensive formulations. Even how to go about choosing γ

remains an open issue. A common approach is to equip the tangent space with an inner product (a Riemannian metric) and then to compute and follow geodesics of $\mathcal{M}$. However, since we could not found any compelling reason for introducing such an artificial structure on our problem we chose a different, simpler approach. The only requirements we have are that the curve passes through x, that is $\gamma(0) = x$, and that it starts off in the direction of ξ, i.e. $\dot{\gamma}(0) = \xi$. In some sense, any curve that satisfies these two conditions will do, and it turns out that with the current problem formulation, finding such curves can be done with very little computational effort, as we will show in the next section.

4. The Manifold of Multiple Homographies.

4.1. The Geometry of the Manifold.

We now return to characterizing the set of valid collections of homographies. As a starting point we use the work of [15]. There it was shown that the set of n-by-m, rank-r matrices forms a $r(n+m-r)$ dimensional manifold embedded in $\mathbb{R}^{nm}$. The tangent space at a point X on this manifold is given by the set of tangent vectors

$$\left\{\Delta = U\Omega + \Gamma V |\ \Gamma \in \mathbb{R}^{n\times r},\ \Omega \in \mathbb{R}^{r\times m},\ \right\}, \qquad (15)$$

where $U \in \mathbb{R}^{n\times r}$ and $U \in \mathbb{R}^{r\times m}$ are the terms in the factorization $X = UV$.

The constraints in (8) clearly identify H as being of at most rank 4, and thus residing on this manifold. In addition, owing to the structure imposed on U in (8), the collections of homographies actually only make up a subset of the rank-4 manifold. Here we give the resulting theorem, summarizing the properties of the homography manifold, along with a brief proof.

Theorem 1. *The set of permissible collections of m homographies of the form given in equation* (8) *is a manifold $\mathcal{H}$, of dimension at most $4m+7$, embedded in $\mathbb{R}^{9m}$.*

The tangent space of $\mathcal{H}$ at a point $H = UV \in \mathcal{H}$ is the set

$$T_{\mathcal{H}}(x) = \left\{\Delta = U\Omega + \Gamma V |\ \Gamma \in \mathcal{L},\ \Omega \in \mathbb{R}^{4\times m}\right\}. \qquad (16)$$

Proof. It follows directly from the restriction of the rank-4 manifold by (8) that $\mathcal{H}$ is a submanifold. The tangent space of $\mathcal{H}$ is then obtained by simply restricting Γ to $\mathcal{L}$ in (15).

A factorization in the form of (8) is not unique since if $H = UV$ then $H = (US)(S^{-1}V)$, where

$$S = \begin{bmatrix} c_1 & 0 & 0 & c_2 \\ 0 & c_1 & 0 & c_3 \\ 0 & 0 & c_1 & c_4 \\ 0 & 0 & 0 & c_5 \end{bmatrix}, \qquad (17)$$

and invertible. It can easily shown that the linear transformation S maps $\mathcal{L}$ onto itself. Hence, $\tilde{U} = (US) \in \mathcal{L}$, $\tilde{V} = (S^{-1}V) \in \mathbb{R}^{4\times m}$ and $\tilde{U}\tilde{V}$ is a factorization that agrees with (8). Consequently, as the dimension of $\mathcal{L}$ is $\dim \mathcal{L} = 3 + 9 = 12$ and S has 5 degrees of freedom, the dimension of $\mathcal{H}$ can be at most $12 + 4m - 5 = 4m + 7$. □

An immediate consequence of this theorem is

Corollary 1. *The formulation in* (8) *provides constraints on as few as only two homographies.*

Proof. With $m = 2$, the manifold $\mathcal{H}$ is of dimension at most $4m + 7 = 15$, embedded in an $9m = 18$ dimensional Euclidean space. □

With theorem 1 we can finally state the problem of finding a valid collection of homographies, that best approximates a set of given homography measurements, as the unconstrained manifold optimization problem

$$\min_{H\in\mathcal{H}} \quad f(H) = ||\tilde{H} - H||. \qquad (18)$$

In order to employ the theory of the previous section we also need to be able to find curves on $\mathcal{H}$ for the purpose of line searches on the manifold. The approach taken here is summarized in the following theorem.

Theorem 2. *A curve on $\mathcal{H}$ passing through the point $H = UV \in \mathcal{H}$ in the direction of ξ is given by*

$$\gamma(t) = H + t\xi + t^2\Lambda \qquad (19)$$

where $\Lambda = \dot{U}\dot{V}$ and $\dot{U} \in \mathcal{L},\ \dot{V} \in \mathbb{R}^{4\times m}$ satisfies

$$\dot{U}V + U\dot{V} = \xi. \qquad (20)$$

Proof. Clearly $\gamma(0) = H$ and $\dot{\gamma}(0) = \xi$. We need only to show that γ maps onto the manifold $\mathcal{H}$. Using (20) we can write

$$\begin{aligned}\gamma(t) = H + t(\dot{U}V + U\dot{V}) + t^2\dot{U}\dot{V} = \\ = (U + \dot{U}t)(V + \dot{V}t)\end{aligned} \qquad (21)$$

Since $\mathcal{L}$ is a linear subspace and both $U, \dot{U} \in \mathcal{L}$, then the first term in (21), $(U + \dot{U}t)$ is in $\mathcal{L}$ for all t and γ is a map onto $\mathcal{H}$. □

This choice of curve representation on $\mathcal{H}$ was motivated by computational requirements. Instead of solving a geodesic equation associated with some choice of some Riemannian metric, as is the common approach, we simply find a solution to the underdetermined linear system in $12 + 4m$ unknowns (20) to obtain γ.

4.2. Optimization on the Manifold of Multiple Homographies.

Before we sum up the previous sections and propose an algorithm, we need to tie in the choice of norm in the objective function of (18). As this choice is highly task specific, we wish to leave this issue as open as possible. Instead we discuss how to implement some commonly used norms. In this work we considered the following norms.

The C-norm, or weighted Frobenius norm, is given by

$$||H||_C^2 = \vec{H}^T C \vec{H} = ||Q\vec{H}||_F^2 \tag{22}$$

where C is a positive-definite weighting matrix and $C = Q^T Q$ its Cholesky factorization. It was indicated in the work by Chen and Suter [2, 3] that such weighted norms can greatly improve de-noising results in the presence of heteroscedastic noise. Note that when C is the identity matrix, this norm becomes the ordinary Frobenius norm.

With this choice of norm, the gradient ∇f and Hessian hess$f$, as defined in (13) and (14), at a point $H = UV$ on the manifold $\mathcal{H}$ becomes

$$\nabla f = -2E(E^T E)^+ E^T C(\vec{\tilde{H}} - \vec{H}) \tag{23}$$

and

$$\text{hess}f = 2E(E^T E)^+ E^T C E (E^T E)^+ E^T, \tag{24}$$

respectively. Here E is the matrix $E = \left[\begin{smallmatrix} U^T \otimes I_9 & I_m \otimes V \end{smallmatrix} \right]$.

Instead of using the Frobenius norm in the right hand side of (22), we can also consider other weighted norms. A popular choice is the Huber norm [9], a twice-differentiable hybrid L_1/L_2 error measure that is more robust to outliers. It is defined as

$$||H||_\mu^2 = \sum_{i=1}^{n} \sigma\left([H]_i\right), \tag{25}$$

$$\sigma(t) = \begin{cases} \frac{1}{2\mu} t^2, & |t| < \mu \\ |t| - \frac{\mu}{2}, & |t| \geq \mu \end{cases}, \tag{26}$$

where μ is a tuning constant, determining the trade-off between L_1 and L_2-norms. By letting $\mu \to 0$ the Huber norm approaches the L_1-norm. Introducing some additional notation

$$W(x) = \text{diag}\left[1 - s_1^2(x),\ ...,\ 1 - s_n^2(x)\right], \tag{27}$$

$$s_i(x) = \begin{cases} -1, & x_i < -\mu \\ 0, & |x_i| \leq \mu \\ 1, & x_i > \mu \end{cases} \tag{28}$$

allows us to write (25) as

$$||H||_\mu^2 = \frac{1}{2\mu} H^T W H + s^T \left[H - \frac{\mu}{2} s \right], \tag{29}$$

$$W = W(\vec{\tilde{H}} - \vec{H}),\ s = s(\vec{\tilde{H}} - \vec{H}). \tag{30}$$

For the Huber-norm, the gradient and Hessian becomes

$$\nabla f = -E(E^T E)^+ E^T \left(\frac{1}{\mu} Q^T W Q (\vec{\tilde{H}} - \vec{H}) + Q^T s \right) \tag{31}$$

and

$$\text{hess}f = \frac{1}{\mu} E(E^T E)^+ E^T Q^T W Q E (E^T E)^+ E^T \tag{32}$$

respectively.

Now we can combine the theory of the previous four sections and finally present our suggested method for minimizing over the set of valid collections of homographies, algorithm 1.

Algorithm 1: Optimization on the Homography Manifold.

input : An initial estimate $H_0 = U_0 V_0 \in \mathcal{H}$.
output: H^* a local minimizer of (18).

```
k = 0;
repeat
    Find ∇f and hessf ;
    Compute the search direction,
    ξ = −(hessf)^{-1} ∇f;
    Determine γ_k by solving U̇_k, V̇_k in (20);
    Search along γ_k for the minima of f(γ_k(t)) ;
    Update, H_{k+1} = γ_k(t*) ;
    k= k+1;
until convergence ;
```

Remarks to Algorithm 1

Input. We currently initialize the algorithm with a random choice of H_0.

3. Depending on the choice of norm ∇f and hess$f$ are given by (23)-(24) or (31)-(32).

4. If C is only positive semidefinite then hess$f$ is not guaranteed to be invertible. In these cases, if necessary use $\xi = -(\text{hess}f + \epsilon I)^{-1} \nabla f$ for some small ϵ ($\epsilon = 1e-3$).

5. The linear system equation (20) is underdetermined, with a whole subspace of solutions. We chose the solution with the smallest norm. It can be shown that this will yield a curve $\gamma_k(t)$ with the smallest curvature. From the discussion in section 3, we thus replace straight line with curves that are as straight as possible.

6. With the curve on the form (19), the function $f(\gamma_k(t))$ is a fourth-degree (piecewise in the case of the Huber-norm) polynomial in t. Minimizing such a one-variable polynomial can be done very efficiently. Here we use the simple Armijo's method.

9. Iteration is terminated when the change in function value is below a given threshold, $|f(H_{k+1}) - f(H_k)| < \epsilon$. We used $\epsilon = 10^{-6}$.

Despite the somewhat complicated definitions and derivations leading up to the above algorithm, it is surprisingly simple. There are no steps in algorithm 1 that are especially demanding. In the experiments carried out in the following section, execution times were on average in the order of tenths of a second for scenes containing a moderate number (< 10) planar surfaces.

It is important to note that the proposed algorithm is a local method, and thus cannot be guaranteed to find the global optimum. However, as we use a line search in conjunction with a descent direction ξ the algorithm is guaranteed to converge to a local minima.

The solution obtained will, as for all local methods, depend on the starting point of this iterative process. How to properly initialize the our algorithm will clearly depend on which objective function is used. We did observe that for the weighted Frobenius norm and the Huber norm our approach is very robust to initialization and that the global minima is indeed obtained in almost all instances, see the following section. This indicates a robustness of the proposed approach and we therefore deliberately left the issue of initialization open to be further addressed in future work.

Additionally, the proposed method displayed locally super-linear convergence rates. However, these observations needs to be theoretically validated.

5. Experiments

A number of problems have been suggested within the literature which require the estimation of self-consistent (constrained) sets of homographies. In [14] the use of such constraints to ensure numerical consistency across views was first mentioned. Other areas of application have included novel view synthesis [1], image sequence alignment and non-rigid motion detection [16].

In this section we first demonstrate the robustness to initialization of the proposed method and then experimentally validate our assertion above that the constraints proposed here represent an improvement over the standard rank-4 constraints. We also show how we can improve on estimating consistent homographies using the weighted Frobenius norm, as well as the Huber norm.

To demonstrate the robustness to initialization of the proposed method we applied algorithm 1 to 300 randomly generated problems, using both the weighted C-norm and the Huber norm. Figure 2 shows a histogram of the distance between the minima obtained by our algorithm and the true minima, for these randomly generated problems. The proposed approach finds the global minima in 97% of the cases.

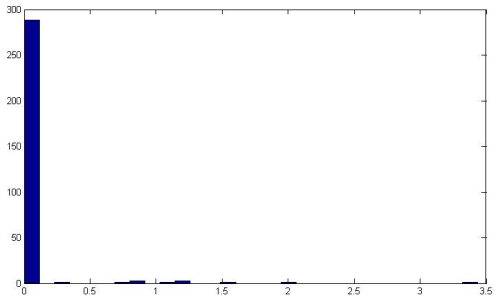


Figure 2. A histogram of the distance between the minima obtained by algorithm 1 and the true minima for 300 randomly generated problems. The proposed algorithm finds the global minima in 97% of the cases.

Next, in order to enable a quantitative assessment of the improvements obtained by exploiting the proposed constraints we evaluated them on a number of synthetically created scene configurations. We randomly generated scenes containing two cameras, five planes and 30 points on each plane. The correspondences, and the plane to which they belong, were assumed to be known for all 150 points. To these ideal points ($\bar{x}_{ij}$ and $\bar{x}'_{ij}$, in the first and second image respectively, with $j = 1...5,\ i = 1...30$) varying levels of Gaussian noise was subsequently added to obtain a collection of corrupted point correspondences, x_{ij} and x'_{ij}. The homography for each plane h_j was then individually estimated from x_{ij} and x'_{ij} using the normalized DLT method. The weighting matrix C was estimated as in [2]. It was the performance of de-noising this set of estimates that was then investigated.

Figure 3 shows how enforcing different constraints improve the estimated homographies for increasing noise levels. The error metric used was the sum of distances between the true image points in the second image $\bar{x}'_{ij}$ and the mapping of the points in the first image $\bar{x}_{ij}$ by homography h_j, and its symmetric counterpart.

$$\varepsilon(H) = \sum_{j=1}^{5}\sum_{i=1}^{30} d(h_j\bar{x}_{ij}, \bar{x}'_{ij}) + d(\bar{x}_{ij}, (h_j)^{-1}\bar{x}'_{ij}) \quad (33)$$

Ideally this distance should be zero.

We then evaluated (33), with no constraints, the rank-4 constraint of [14] and the proposed manifold constraints.

For comparison we also include the results from using bundle adjustment to minimize the Gold standard error,

$$\min_{A,a,v_i,\hat{x}_{ij}} \quad \sum_{j=1}^{5}\sum_{i=1}^{30} d(\hat{x}_{ij}, x_{ij})^2 + \\ + d(H(A,a,v_i)\hat{x}_{ij}, x'_{ij})^2. \quad (34)$$

The results in figure 3 were obtained by averaging a large number of simulations. It can clearly be seen that, enforc-

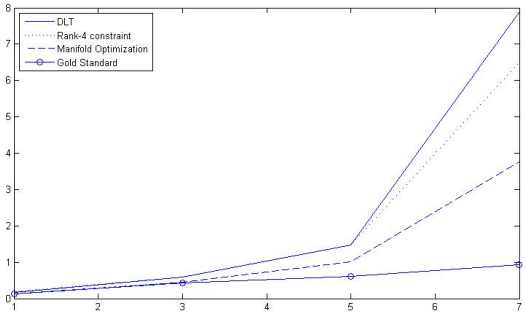


Figure 3. The averaged error $\varepsilon(H)$ as a function of increasing noise levels (standard deviation in pixels), for different constraints.

Method	Time
Normalized DLT	0.027s
Truncated SVD	0.0034s
Manifold Optimization	0.12s
Bundle Adjustment	52s

Table 1. Computational time for different methods.

ing the additional constraints in (8) greatly improves the estimated homographies over the rank-4 constraint. But, as expected, it is the Gold standard approach that produce the best results. However, this is a formulation that will only be applicable to L_2 norms, whereas in our suggested approach one can choose any differentiable norm. The Gold standard method will also result in a much larger optimization problem, involving more than 300 variables, thus arguably having higher computational demands and being more sensitive to initialization. We include table 1 as an indication of the time requirements for the different methods used in this section. The software used was in all instances standard Matlab implementations. There are sparsity issues that should be taken into consideration as well, in the Gold standard as well as in our proposed method. However, as these still remain to be investigated in the latter case, we chose to compare the time requirements for dense implementations only. The gold standard method, even when initialized with the ideal values $\bar{x}$ and $\bar{x}'$ is it extremely time consuming.

Next, we illustrate our method on a real world situation. We again assume that the correspondences, and the plane to which they belong, were known for a number of image points in the scene. If one or more estimated homographies in a scene has been corrupted by noise the homography constraints derived above will not be fulfilled. It is then possible to de-noise the estimates by enforcing the constraints through the minimization of (18). An example of this process can be seen in figure 4. Here we have added additional noise to the measured image points on the right-most planar surfaces in the image pair. Using the same approach as above we can recover a more accurate estimate as well as consistent set of homographies compared to estimating the homographies individually.

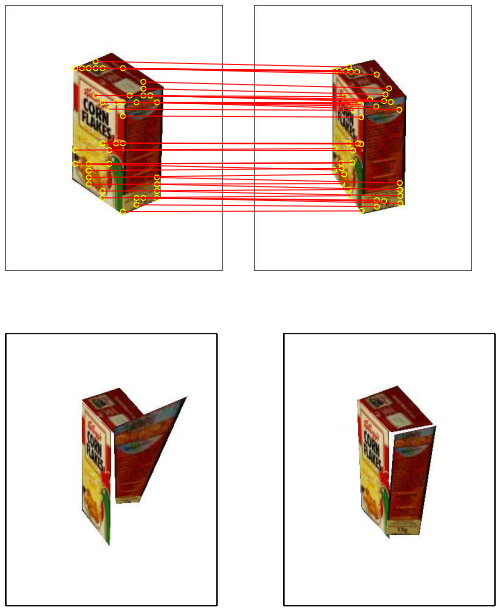
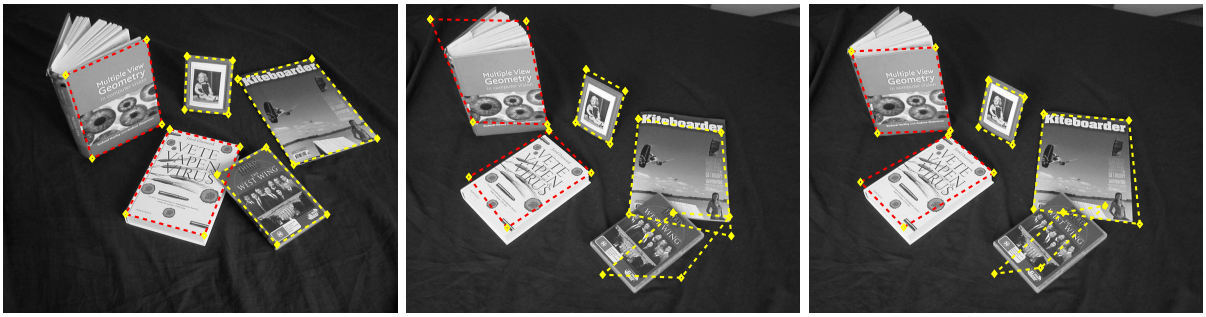


Figure 4. Original image pair (top) used for homography estimation. If one of the estimates have been corrupted by noise (bottom left) and improved estimate can be achieved (bottom right) by enforcing the constraints in (8).

Finally, we illustrate the potential of the robust Huber-norm. If one assumes that the individual homographies has been well estimated, then manifold constraints can only be violated if the underlying motion is not rigid. We know that the Huber-norm is an error measure that is robust to outliers. If we view the homography corresponding to the non-rigid motion as a vector of outliers and employ the Huber-norm, then the remaining homographies should still be estimated well and the outlier homography should be the only one with a large error. This is typically not the case when using L_2 type norms.

An example can be seen in figure 5, a simple scene consisting of five different planes, where the planar surface in the bottom right hand corner has moved relative to the rest of the scene. The result of enforcing the homography constraints using the C-norm and Huber-norm respectively is shown in figure 5 (b)-(c). Clearly, the latter choice of norm produces reasonable estimates for the four rigidly moving homographies but not for the fifth. As opposed to the C-norm which yield large errors for three out of the five planes in the scene.



(a) (b) (c)

Figure 5. An image pair, (a) and (b)-(c), consisting of five different planes. The resulting recovered homographies, after enforcing constraints $H \in \mathcal{H}$, using the C-norm (b) and the Huber-norm (c).

6. Conclusions

In this paper we have provided a new take on the geometric relationship between the image motion of multiple planar surfaces. Our approach constitutes a reformulation of previously known constraints and is based on optimization on manifolds.

The are two main contributions. Firstly, the novel characterization of the set of valid collections of homographies. We have shown that the previously known rank-4 constraint on multiple homographies can be improved. It was proven that this set of homographies actually make up a low-dimensional differentiable manifold embedded in Euclidean space. Secondly, we proposed a computationally efficient algorithm, a generalization of Newton's method, for minimizing any twice differentiable objective function on this homography manifold.

The strength of these additional constraints, as well as the efficiency of the suggested optimization method, was validated experimentally on both synthetic and real world data.

References

[1] S. Avidan and A. Shashua. Novel view synthesis in tensor space. In *In Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1034–1040, 1997.

[2] P. Chen and D. Suter. Homography estimation and heteroscedastic noise–a first order perturbation analysis. *Technical Report MECSE-32-2005, Monash University, Australia*, 2005.

[3] P. Chen and D. Suter. An analysis of linear subspace approaches for computer vision and pattern recognition. *Int. J. Comput. Vision*, 68(1):83–106, 2006.

[4] P. Chen and D. Suter. Rank constraints for homographies over two views: Revisiting the rank four constraint. *Int. Journal of Computer Vision*, To Appear.

[5] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20(2):303–353, 1999.

[6] O. Faugeras. *Three-dimensional computer vision: a geometric viewpoint*. MIT Press, Cambridge, MA, USA, 1993.

[7] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[8] U. Helmke, K. Hüper, P. Y. Lee, and J. Moore. Essential matrix estimation using gauss-newton iterations on a manifold. *Int. J. Comput. Vision*, 74(2):117–136, 2007.

[9] P. J. Huber. Robust regression: Asymptotics, conjectures, and monte carlo. In *Annals of Statistics*, volume 1, pages 799–821, 1973.

[10] K. Huper and J. Trumpf. Newton-like methods for numerical optimization on manifolds. In *In Proc. of the 38th Asilomar Conference on Signals, Systems and Computers*, pages 136–139, 2004.

[11] F. Kahl, S. Agarwal, M. Chandraker, D. Kriegman, and S. Belongie. Practical global optimization for multiview geometry. *International Journal of Computer Vision*, 79(3):271–284, 2008.

[12] F. Kahl and R. Hartley. Multiple view geometry under the l-infinity norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(9):1603–1617, 2008.

[13] E. Malis and R. Cipolla. Multi-view constraints between collineations: Application to self-calibration from unknown planar structures. In *ECCV '00: Proceedings of the 6th European Conference on Computer Vision-Part II*, pages 610–624, London, UK, 2000. Springer-Verlag.

[14] A. Shashua and S. Avidan. The rank 4 constraint in multiple (≥ 3) view geometry. In *ECCV 1996: Proceedings of the 4th European Conference on Computer Vision-Volume II*, pages 196–206, London, UK, 1996. Springer-Verlag.

[15] H. U. and M. J.B. *Optimization and Dynamical Systems*. Springer-Verlag, 1993.

[16] L. Zelnik-Manor and M. Irani. Multiview constraints on homographies. In *IEEE Trans. Pattern Anal. Mach. Intell.*, volume 24, pages 214–223, Washington, DC, USA, 2002. IEEE Computer Society.