Uncertainty Modeling for Optimal Structure from Motion

Daniel D. Morris[1], Kenichi Kanatani[2], and Takeo Kanade[1]

[1] Robotics Institute, Carnegie Mellon University,
Pittsburgh, PA 15213, USA
{ddmorris, tk}@ri.cmu.edu
[2] Department of Computer Science, Gunma University,
Kiryu, Gunma 376-8515, Japan
kanatani@cs.gunma-u.ac.jp

Abstract. The parameters estimated by Structure from Motion (SFM) contain inherent indeterminacies which we call gauge freedoms. Under a perspective camera, shape and motion parameters are only recovered up to an unknown similarity transformation. In this paper we investigate how covariance-based uncertainty can be represented under these gauge freedoms. Past work on uncertainty modeling has implicitly imposed gauge constraints on the solution before considering covariance estimation. Here we examine the effect of selecting a particular gauge on the uncertainty of parameters. We show potentially dramatic effects of gauge choice on parameter uncertainties. However the inherent geometric uncertainty remains the same irrespective of gauge choice. We derive a Geometric Equivalence Relationship with which covariances under different parametrizations and gauges can be compared, based on their true geometric uncertainty. We show that the uncertainty of gauge invariants exactly captures the geometric uncertainty of the solution, and hence provides useful measures for evaluating the uncertainty of the solution. Finally we propose a fast method for covariance estimation and show its correctness using the Geometric Equivalence Relationship.

1 Introduction

It is well known that, for accurate 3D reconstruction from image sequences, statistically optimal results are obtained by bundle adjustment [2,3,5,6,13,16]. This is just Maximum Likelihood estimation for independent, isotropic Gaussian noise, and is also used by photogrammetrists. Current research generally focuses on two areas: (1) simplicity of solution, which includes finding a closed form approximate solutions such as the Factorization method [4,8–12], and (2) efficiency, which includes finding fast or robust numerical schemes [1,2].

An important third area to address is the quantitative assessment of the reliability of the solution. While some work has incorporated uncertainty analyzes of the results [9,14–16], none has investigated the effect of parameter indeterminacies on the uncertainty modeling. These indeterminacies are inherent to SFM

and have a significant effect on parameter uncertainties. Our goal is to create a framework for describing the uncertainties and indeterminacies of parameters used in Structure from Motion (SFM). We can then determine how both these uncertainties and indeterminacies affect the real geometric measurements recovered by SFM.

The standard measure for uncertainty is the covariance matrix. However in SFM there is a uniqueness problem for the solution and its variance due to inherent indeterminacies: the estimated object feature positions and motions are only determined up to a overall translation, rotation and scaling. Constraining these global quantities we call choosing a *gauge*. Typically a covariance matrix describes the second order moments of a perturbation around a unique solution. In past work [9, 15, 16] indeterminacies are removed by choosing an arbitrary gauge, and then the optimization is performed under these gauge constraints and the recovered shape and motion parameters along with their variances are expressed in this gauge.

In this paper we provide an analysis of the effects of indeterminacies and gauges on covariance-based uncertainty models. While the choice of gauge can dramatically affect the magnitude and values in a covariance matrix, we show that these effects are superficial and the underlying geometric uncertainty is unaffected. To show this we derive a Geometric Equivalence Relationship between the covariance matrices of the parameters that depends only on the essential geometric component in the covariances. Hence we are able to propose a covariance-based description of parameter uncertainties that does not require gauge constraints. Furthermore we show how this parametric uncertainty model can be then used to obtain an uncertainty model for actual geometric properties of the shape and motion which are gauge-invariant. Optimization is achieved in an efficient free-gauge manner and we propose a fast method for obtaining covariance estimates when there are indeterminacies.

2 Geometric Modeling

2.1 Camera Equations

Here we describe an object and camera system in a camera–centered coordinate system. Analogous equations could be derived in other coordinate systems. Suppose we track N rigidly moving feature points $\mathbf{p}_\alpha$, $\alpha = 1, \ldots, N$, in M images. Let $\mathbf{p}_{\kappa\alpha}$ be the 2–element image coordinates of $\mathbf{p}_\alpha$ in the κth image. We identify the camera coordinate system with the XYZ world coordinate system, and choose an object coordinate system in the object. Let $\mathbf{t}_\kappa$ be the origin of the object coordinate system in the κ'th image, $\mathbf{R}_\kappa$ be a 3×3 rotation matrix which specifies its orientation and $\mathbf{s}_\alpha$ be the coordinates of the feature point, $\mathbf{p}_\alpha$, in the object coordinate system. Thus the position of feature point $\mathbf{p}_\alpha$ with respect to the camera coordinate system in the κth image is $\mathbf{R}_\kappa \mathbf{s}_\alpha + \mathbf{t}_\kappa$.

Assume we have a projection operator $\Pi : \mathcal{R}^3 \rightarrow \mathcal{R}^2$ which projects a point in 3D to the 2D image plane. We can then express the image coordinates, of

feature $\mathbf{p}_\alpha$ as:

$$\mathbf{p}_{\kappa\alpha} = \Pi[\mathbf{K}_\kappa(\mathbf{R}_\kappa \mathbf{s}_\alpha + \mathbf{t}_\kappa)] \tag{1}$$

where $\mathbf{K}_\kappa$ is a 3×3 internal camera parameter matrix [2] containing quantities such as focal length for each image. While these parameters can be estimated along with shape and motion parameters, for simplicity we ignore them in the rest of the paper and assume $\mathbf{K}_\kappa$ is just the identity matrix for orthography and $diag([f, f, 1])$ for perspective projection with focal length f. Various camera models can be defined by specifying the action of this projection operator on a vector $(X, Y, Z)^\top$. For example we define the projection operators for orthography and perspective projection respectively in the following way:

$$\Pi_o[\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}] = \begin{pmatrix} X \\ Y \end{pmatrix}, \quad \Pi_p[\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}] = \begin{pmatrix} X/Z \\ Y/Z \end{pmatrix}, \tag{2}$$

Equation (1) can be applied to all features in all images, and then combined in the form:

$$\mathbf{p} = \mathbf{\Pi}(\boldsymbol{\theta}) \tag{3}$$

where $\mathbf{p} = (\mathbf{p}_{11}^\top, \mathbf{p}_{12}^\top, \mathbf{p}_{13}^\top, \ldots, \mathbf{p}_{MN}^\top)^\top$ is a vector containing all the image feature coordinates in all images, and $\boldsymbol{\theta}$ is a vector containing the shape and motion parameters, $\mathbf{R}_\kappa$, $\mathbf{s}_\alpha$, $\mathbf{t}_\kappa$, and possibly unknown internal camera parameters, for all object features and images, and $\mathbf{\Pi}$ is the appropriate combination of the projection matrices. More details can be found in [7].

2.2 Parameter Constraints

Not all of the parameters in $\boldsymbol{\theta}$ are independent and some need to be constrained. In particular the columns of each rotation matrix, $\mathbf{R}_\kappa$, must remain unit orthogonal vectors. Small perturbations of rotations are parametrized by a 3-vector: $\Delta\boldsymbol{\Omega}_\kappa$ which to first order maintain the rotation properties [3]. Let $\mathcal{T}$ be the manifold of valid vectors $\boldsymbol{\theta}$ such that all solutions for $\boldsymbol{\theta}$ lie in $\mathcal{T}$. $\mathcal{T}$ will be a manifold of dimension n, where n is the number of parameters needed to locally specify the shape and motion, 3 for each rotation, 3 for each translation, and 3 for each 3D feature point, plus any internal camera parameters that must be estimated. So in general for just motion and shape, the number of unknown parameters is: $n = 3N + 6M$.

2.3 Indeterminacies

The camera equations (1) and (3) contain a number of indeterminacies. There are two reasons for these indeterminacies: first the object coordinate system can be selected arbitrarily, and second the projection model maps many 3D points to a single 2D point. These are specified as follows:

Coordinate System Indeterminacies

If we rotate and then translate the coordinate system by $\mathbf{R}$ and $\mathbf{t}$ respectively, we obtain the following transformed shape and motion parameters:

$$\mathbf{s}'_\alpha = \mathbf{R}^\top(\mathbf{s}_\alpha - \mathbf{t}),\ \ \mathbf{R}'_\kappa = \mathbf{R}_\kappa \mathbf{R},\ \ \mathbf{t}'_\kappa = \mathbf{R}_\kappa \mathbf{t} + \mathbf{t}_\kappa\,. \tag{4}$$

We note that $\mathbf{R}'_\kappa \mathbf{s}'_\alpha + \mathbf{t}'_\kappa = \mathbf{R}_\kappa \mathbf{s}_\alpha + \mathbf{t}_\kappa$, and hence irrespective of the projection model, equations (1) and (3) must be ambiguous to changes in coordinates.

Projection Indeterminacies

Many different geometric solutions project onto the same points in the image. In orthography the depth or Z component does not affect the image, and hence the projection is invariant to the transformation:

$$\mathbf{t}'_\kappa = \mathbf{t}_\kappa + d_\kappa \mathbf{k} \tag{5}$$

for any value d_κ. Orthography has a discrete reflection ambiguity, but since it is not differential we do not consider it. Perspective projection has a scale ambiguity such that if we transform the shape and translation by a scale s:

$$\mathbf{s}'_\alpha = s\mathbf{s}_\alpha\,,\ \text{ and }\ \mathbf{t}'_\kappa = s\mathbf{t}_\kappa\,, \tag{6}$$

we find that $\Pi_p[\mathbf{K}_\kappa(\mathbf{R}_\kappa \mathbf{s}'_\alpha + \mathbf{t}'_\kappa)] = \Pi_p[\mathbf{K}_\kappa(\mathbf{R}_\kappa \mathbf{s}_\alpha + \mathbf{t}_\kappa)]$.

2.4 Solution Manifold

Since the camera equations contain these indeterminacies, then given the measurement data, $\mathbf{p}$, there is not a unique shape and motion parameter set, $\boldsymbol{\theta}$, that maps to this. Rather equation (3) is satisfied by a manifold, $\mathcal{M}$, of valid solutions within $\mathcal{T}$ which are all mapped to the same $\mathbf{p}$. This manifold has dimension, r, given by the number of infinitesimal degrees of freedom at a given point. From the ambiguity equations (4-6) we obtain $r = 7$ for perspective projection and $r = M + 6$ under orthography. Figure 1 illustrates a solution $\boldsymbol{\theta} \in \mathcal{M}$.

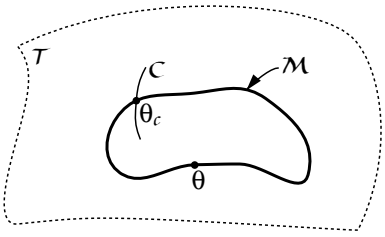

Fig. 1. An illustration of a curve representing a manifold $\mathcal{M}$ of solution vectors $\boldsymbol{\theta}$, all lying in the parameter space $\mathcal{T}$. Choosing a gauge $\mathcal{C}$, which intersects the manifold at one point, defines a unique solution $\boldsymbol{\theta}_\mathcal{C}$.

2.5 Gauge Constraints

In order to remove the ambiguity of the solution we can define a gauge or manifold of points: $\mathcal{C}$. Let $\mathcal{C}$ contain all those points in $\mathcal{T}$ that satisfy a set of r constraint equations:

$$c_i(\boldsymbol{\theta}) = 0 \;\text{ for }\; 1 \leq i \leq r. \tag{7}$$

The gauge $\mathcal{C}$ will thus have dimension $n - r$. We require that $\mathcal{C}$ intersect $\mathcal{M}$ transversally and at most at one point per connected component of $\mathcal{M}$. The intersection of $\mathcal{C}$ and $\mathcal{M}$ thus provides unique solution within a connected component of $\mathcal{M}$, as illustrated in Figure 1. However there may be ambiguities between components of $\mathcal{M}$, such as the reflection ambiguity in orthography.

For example, we could define an arbitrary gauge with the following constraints:

$$\sum_{\alpha=1}^{N} \mathbf{s}_\alpha = \mathbf{0}, \;\; \mathbf{R}_1 = \mathbf{I}, \;\; \sum_{\alpha=1}^{N} \mathbf{s}_\alpha^\top \mathbf{s}_\alpha = 1. \tag{8}$$

This fixes the origin of the object coordinate system in its centroid, aligns the object coordinate system with the first image, and fixes the scale. In orthography the scale constraint is omitted, but we add the constraint set: $t_{z_\kappa} = 0$ on the Z component of translation.

We note that this, or any other choice of gauge is arbitrary, and does not affect the geometry. It does affect our parameter estimates and their uncertainties, but in ways that do not affect the geometric meaning of the results.

3 Uncertainty in Data Fitting

When there is noise in the measured data, there will be a resulting uncertainty in the recovered parameters, which we would like to represent by a covariance matrix. However, when indeterminacies exist, the solution will be a manifold rather than a point, and standard perturbation analysis cannot be performed. The usual approach, in dealing with this, is to choose a gauge and constrain the solution to lie in this gauge. While this approach is a valid, it introduces additional constraints into the estimation process, and the resulting uncertainty values are strongly dependent on the choice of gauge. In this section we ask the question: How can we estimate the geometric uncertainty without depending on an arbitrary selection of a gauge? To answer this we introduce gauge invariants whose uncertainty does not depend on gauge choice. We also derive a Geometric Equivalence Relationship that considers only this "true" geometric uncertainty. Along the way we derive the *normal form* for the covariance which gives us a convenient way to calculate uncertainty without having to explicitly specify a gauge.

3.1 Perturbation Analysis

First we derive an uncertainty measure in an arbitrary gauge. We assume that the noise is small, and thus that the first order terms dominate. When the noise is Gaussian the first order terms exactly describe the noise. The measured data, $\mathbf{p}$ is a result of the true feature positions, $\bar{\mathbf{p}}$, corrupted by noise, $\Delta\mathbf{p}$:

$$\mathbf{p} = \bar{\mathbf{p}} + \Delta\mathbf{p}. \tag{9}$$

The noise $\Delta\mathbf{p}$ is a random variable of the most general type, not necessarily independent for different points, but it is assumed to have zero mean and known variance[1]:

$$V_p[\mathbf{p}] = E\{\Delta\mathbf{p}\Delta\mathbf{p}^\top\}. \tag{10}$$

We note that in the special case when feature points are independent, $V[\mathbf{p}]$ will be block diagonal with the 2×2 block diagonal elements giving the independent feature covariances.

Given this uncertainty in the measured data, let $\hat{\boldsymbol{\theta}}$ be our estimator of the shape and motion parameters. There is no unique true solution unless we restrict our estimation to a particular gauge. If we choose gauge $\mathcal{C}$ our estimator can be written as: $\hat{\boldsymbol{\theta}}_\mathcal{C} = \bar{\boldsymbol{\theta}}_\mathcal{C} + \Delta\boldsymbol{\theta}_\mathcal{C}$, for true solution $\bar{\boldsymbol{\theta}}_\mathcal{C}$ and perturbation $\Delta\boldsymbol{\theta}_\mathcal{C}$. The perturbation $\Delta\boldsymbol{\theta}_\mathcal{C}$ and its variance, $V[\Delta\boldsymbol{\theta}_\mathcal{C}]$, both lie in the tangent plane to the gauge manifold, $T_{\bar{\boldsymbol{\theta}}_\mathcal{C}}[\mathcal{C}]$.

We expand equation (3) around $\bar{\boldsymbol{\theta}}_\mathcal{C}$ and get to first order:

$$\nabla^\mathcal{T}_{\boldsymbol{\theta}}\mathbf{\Pi}(\bar{\boldsymbol{\theta}}_\mathcal{C})\Delta\boldsymbol{\theta}_\mathcal{C} = \Delta\mathbf{p} \tag{11}$$

where $\nabla^\mathcal{T}_{\boldsymbol{\theta}}$ is the gradient with respect to $\boldsymbol{\theta}$ in the manifold $\mathcal{T}$. We then split the perturbations, $\Delta\boldsymbol{\theta}_\mathcal{C}$ into two components, those in $T_{\boldsymbol{\theta}_\mathcal{C}}[\mathcal{M}]$ and those in $T_{\boldsymbol{\theta}_\mathcal{C}}[\mathcal{M}]^\perp$ as shown in Figure 2:

$$\Delta\boldsymbol{\theta}_\mathcal{C} = \Delta\boldsymbol{\theta}_\mathcal{C}{}^{\parallel\mathcal{M}} + \Delta\boldsymbol{\theta}_\mathcal{C}{}^{\perp\mathcal{M}}. \tag{12}$$

The gradient $\nabla^\mathcal{T}_{\boldsymbol{\theta}}\mathbf{\Pi}(\bar{\boldsymbol{\theta}})$ is orthogonal to the tangent plane of $\mathcal{M}$ and has rank $n-r$. We can thus solve for the orthogonal perturbations:

$$\Delta\boldsymbol{\theta}_\mathcal{C}{}^{\perp\mathcal{M}} = (\nabla^\mathcal{T}_{\boldsymbol{\theta}}\mathbf{\Pi}(\bar{\boldsymbol{\theta}}_\mathcal{C}))^-_{n-r}\Delta\mathbf{p}, \tag{13}$$

where "$-$" denotes the Moore-Penrose generalized inverse[2] constrained to have rank $n-r$. We call the covariance of this orthogonal component the *normal covariance*:

$$V_{\perp\mathcal{M}}[\boldsymbol{\theta}] = (\nabla^\mathcal{T}_{\boldsymbol{\theta}}\mathbf{\Pi}(\bar{\boldsymbol{\theta}}))^-_{n-r}V_p(\nabla^\mathcal{T}_{\boldsymbol{\theta}}\mathbf{\Pi}(\bar{\boldsymbol{\theta}}))^{-\top}_{n-r}. \tag{14}$$

The normal covariance is expressed at a particular solution, $\boldsymbol{\theta}$, and depends on our choice of parametrization and implicitly assumes a metric over parameter

[1] We can extend this to the case when variance is known only up to a scale factor

[2] The Moore-Penrose generalized inverse is defined such that if $A = U\Lambda V^\top$ by SVD, then $A^-_N = V\Lambda^-_N U^\top$, where Λ^-_N has the first N singular values inverted on the diagonal, and the rest zeroed.

space. But it does not require explicit gauge constraints, (rather implicitly assumes a gauge normal to the manifold), and as we shall see, it incorporates all of the essential geometric uncertainty in the solution.

When the indeterminacies are removed by adding constraints the normal covariance must be obliquely projected onto the appropriate constraint surface. The uncertainty in the gauge will be in its tangent plane: $\Delta\boldsymbol{\theta}_{\mathcal{C}} \in T[\mathcal{C}]$. We already know the perturbation, $\Delta\boldsymbol{\theta}_{\mathcal{C}}{}^{\perp\mathcal{M}}$, orthogonal to $T[\mathcal{M}]$, and so it only remains to derive the component parallel to $T[\mathcal{M}]$ as illustrated in Figure 2.

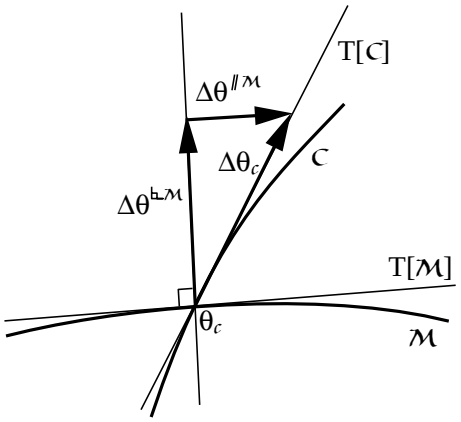

Fig. 2. An illustration of the oblique projection of perturbations along the solution tangent space, $T[\mathcal{M}]$, and onto the gauge manifold tangent space $T[\mathcal{C}]$: $\Delta\boldsymbol{\theta}_{\mathcal{C}} = \Delta\boldsymbol{\theta}_{\mathcal{C}}{}^{\perp\mathcal{M}} + \Delta\boldsymbol{\theta}_{\mathcal{C}}{}^{\parallel\mathcal{M}}$. This projection transforms the normal covariance matrix into the local gauge covariance.

Let U be a matrix with r columns spanning $T[\mathcal{M}]$ at $\boldsymbol{\theta}_{\mathcal{C}}$, and let V be a matrix with r columns spanning the space orthogonal to $T[\mathcal{C}]$ at $\boldsymbol{\theta}_{\mathcal{C}}$. Then we can express equation (12) as:

$$\Delta\boldsymbol{\theta}_{\mathcal{C}} = \Delta\boldsymbol{\theta}_{\mathcal{C}}{}^{\perp\mathcal{M}} + U\mathbf{x}. \tag{15}$$

for some unknown coefficients $\mathbf{x}$. The fact that this perturbation is in the constraint tangent plane, implies that $V^{\top}\Delta\boldsymbol{\theta}_{\mathcal{C}} = 0$. Applying this to (15) and eliminating $\mathbf{x}$ we obtain:

$$\Delta\boldsymbol{\theta}_{\mathcal{C}} = Q^{\mathcal{C}}\Delta\boldsymbol{\theta}_{\mathcal{C}}{}^{\perp\mathcal{M}} \tag{16}$$

where $Q^{\mathcal{C}} = \mathbf{I} - U(V^{\top}U)^{-1}V^{\top}$ is our oblique projection operator along $T[\mathcal{M}]$. The covariance of $\boldsymbol{\theta}$ in this gauge is then given by:

$$V_{\mathcal{C}}[\boldsymbol{\theta}_{\mathcal{C}}] = Q^{\mathcal{C}}V_{\perp\mathcal{M}}[\boldsymbol{\theta}_{\mathcal{C}}]Q^{\mathcal{C}\top}. \tag{17}$$

3.2 Inherent Geometric Uncertainty

The camera equations provide geometric constraints on the measurements. Parameters containing indeterminacies correspond to entities not fully constrained

by the camera equations, whereas parameters which have a unique value over the solution manifold are fully constrained. These fully constrained parameters describe the "true" geometric entities. They can be uniquely recovered, (up to possibly a discrete ambiguity), from the camera equations. Having a unique value on the solution manifold means that the parameter is invariant to gauge transformations on the solution. We call these gauge invariants.

Not only are the values of gauge invariants unique, but given the covariance of the measured data, the covariance of the invariant is uniquely obtainable. However, the covariance of parameters containing indeterminacies will not be uniquely specified and many possible "geometrically equivalent" covariances can be obtained that correspond to the same measurement covariance. In this section we derive a Geometric Equivalence Relationship for parameters that contain indeterminacies. This permits us to test whether covariances of these parameters under different gauges correspond to the same underlying measurement covariance or not. Finally we propose a fast method for covariance estimation and show its correctness using the Geometric Equivalence Relationship.

Let us assume that we are measuring an invariant property, $I(\boldsymbol{\theta})$, of the solution. Consider the estimators in two gauges: $\boldsymbol{\theta}_{\mathcal{C}}$ and $\boldsymbol{\theta}_{\mathcal{C}'}$ with uncertainties: $\Delta\boldsymbol{\theta}_{\mathcal{C}}$ and $\Delta\boldsymbol{\theta}_{\mathcal{C}'}$ in their corresponding tangent planes. Let $\frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}$ be the Jacobian matrix that maps perturbations in the tangent plane of $\mathcal{C}'$ to perturbations in the tangent plane of $\mathcal{C}$:

$$\Delta\boldsymbol{\theta}_{\mathcal{C}} = \frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}\Delta\boldsymbol{\theta}_{\mathcal{C}'}. \tag{18}$$

The invariant property will have the same value for both solutions: $I(\boldsymbol{\theta}_{\mathcal{C}}) = I(\boldsymbol{\theta}_{\mathcal{C}'})$. Moreover, since I is invariant to all points in $\mathcal{M}$, it must also be invariant to infinitesimal perturbations in $\mathcal{M}$, and hence its gradient must be orthogonal to the tangent plane of $\mathcal{M}$:

$$\nabla^{\mathcal{T}}_{\boldsymbol{\theta}} I \in T[\mathcal{M}]^{\perp}. \tag{19}$$

A perturbation of the invariant at $\boldsymbol{\theta}_{\mathcal{C}}$ can be written:

$$\Delta I(\boldsymbol{\theta}_{\mathcal{C}}) = \nabla^{\mathcal{T}}_{\boldsymbol{\theta}} I\Delta\boldsymbol{\theta}_{\mathcal{C}} = \nabla^{\mathcal{T}}_{\boldsymbol{\theta}} I\frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}\Delta\boldsymbol{\theta}_{\mathcal{C}'}. \tag{20}$$

The variance of the invariant can be calculated using both components of this equation:

$$V[I] = \nabla^{\mathcal{T}}_{\boldsymbol{\theta}} IV[\boldsymbol{\theta}_{\mathcal{C}}]\nabla^{\mathcal{T}}_{\boldsymbol{\theta}} I^{\top} = \nabla^{\mathcal{T}}_{\boldsymbol{\theta}} I\frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}V[\boldsymbol{\theta}_{\mathcal{C}'}]\frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}^{\top}\nabla^{\mathcal{T}}_{\boldsymbol{\theta}} I^{\top}. \tag{21}$$

The covariances of parameters with indeterminacies may have "non-geometric" components along the tangent plane of the solution manifold. This equation transforms these covariances into the uniquely defined covariance of a gauge invariant.

We then apply the orthogonal constraint from equation (19) to both expressions for $V[I]$ and obtain the following result:

$$\mathbf{u}^{\top}(V[\boldsymbol{\theta}_{\mathcal{C}}] - \frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}V[\boldsymbol{\theta}_{\mathcal{C}'}]\frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}^{\top})\mathbf{u} = 0, \quad \forall\mathbf{u} \in T_{\boldsymbol{\theta}_{\mathcal{C}}}[\mathcal{M}]^{\perp}. \tag{22}$$

This means that the difference between the covariance and the transformed covariance: $V[\boldsymbol{\theta}_{\mathcal{C}}] - \frac{\partial \boldsymbol{\theta}_{\mathcal{C}}}{\partial \boldsymbol{\theta}_{\mathcal{C}'}} V[\boldsymbol{\theta}_{\mathcal{C}'}] \frac{\partial \boldsymbol{\theta}_{\mathcal{C}}}{\partial \boldsymbol{\theta}_{\mathcal{C}'}}^{\top}$ must lie in the the tangent space $T_{\boldsymbol{\theta}_{\mathcal{C}}}[\mathcal{M}]$. Or equivalently we can say that these two variances have the same orthogonal component to $T[\mathcal{M}]$ at $\boldsymbol{\theta}_{\mathcal{C}}$. We denote this relationship as: $V[\boldsymbol{\theta}_{\mathcal{C}}] \equiv V[\boldsymbol{\theta}_{\mathcal{C}'}] \bmod \mathcal{M}$. Thus we have:

Geometric Equivalence Relationship *The covariance matrices* $V[\boldsymbol{\theta}_{\mathcal{C}}]$ *and* $V[\boldsymbol{\theta}_{\mathcal{C}'}]$ *are geometrically equivalent if and only if*

$$V[\boldsymbol{\theta}_{\mathcal{C}}] \equiv V[\boldsymbol{\theta}_{\mathcal{C}'}] \bmod \mathcal{M}. \tag{23}$$

In essence this says that at a point $\boldsymbol{\theta} \in \mathcal{M}$, it is only the component of the covariance that is not in the tangent plane that contributes to the geometric uncertainty. Any matrix satisfying this equivalence relationship captures the geometric uncertainty of the parameters. The normal form of the covariance calculated from equation (14) is a natural choice that captures this uncertainty for a given parametrization, and does not require constraints to be specified. From this relationship we see that the covariance in any gauge is equivalent to the normal covariance, i.e.: $V_{\mathcal{C}}[\boldsymbol{\theta}_{\mathcal{C}}] \equiv V_{\perp \mathcal{M}}[\boldsymbol{\theta}] \bmod \mathcal{M}$. Thus the covariance of an invariant can be calculated directly from either of these covariances by transforming them with the invariant gradient, $\nabla_{\boldsymbol{\theta}}^{\mathcal{T}} I$, as in equation (21).

4 Maximum Likelihood Estimation

It is known that *Maximum Likelihood* (ML) estimation is unbiased and obtains the optimal shape and motion parameters. The ML solution is obtained by minimizing the cost:

$$J = (\mathbf{p} - \boldsymbol{\Pi}(\boldsymbol{\theta}))^{\top} V_p^{-1} (\mathbf{p} - \boldsymbol{\Pi}(\boldsymbol{\theta}))). \tag{24}$$

where $\boldsymbol{\theta} \in \mathcal{T}$. The minimum value of this will have the same camera indeterminacies described in section 2.3, and hence determine a manifold, $\mathcal{M}$, of geometrically equivalent solutions. A unique solution can be obtained by choosing an arbitrary gauge $\mathcal{C}$.

4.1 Free-Gauge Solution

Instead of constraining our minimization process with our chosen gauge $\mathcal{C}$, at each step we would like to choose a gauge orthogonal to the solution manifold $\mathcal{M}$, and proceed in that direction. We expect this to give better convergence to the manifold $\mathcal{M}$ especially when our desired gauge $\mathcal{C}$ has a large oblique angle to $\mathcal{M}$. Once any point on $\mathcal{M}$ is achieved, it is easy to transform this solution into any desired gauge.

Levenberg-Marquardt (LM) minimization is a combination of Gauss-Newton and gradient descent. The gradient of J is obtained as

$$\nabla_{\boldsymbol{\theta}} J = -2 \nabla_{\boldsymbol{\theta}}^{\mathcal{T}} \boldsymbol{\Pi}(\boldsymbol{\theta}) V_p^{-1} (\mathbf{p} - \boldsymbol{\Pi}(\boldsymbol{\theta})), \tag{25}$$

and the Gauss-Newton approximation for the Hessian:

$$\nabla^2_{\boldsymbol{\theta}} J \approx \frac{1}{2} E\{\nabla_{\boldsymbol{\theta}} J \nabla_{\boldsymbol{\theta}} J^\top\} = 2 \nabla^{\mathcal{T}}_{\boldsymbol{\theta}} \mathbf{\Pi}(\boldsymbol{\theta}) V_p^{-1} \nabla^{\mathcal{T}}_{\boldsymbol{\theta}} \mathbf{\Pi}(\boldsymbol{\theta})^\top. \tag{26}$$

Gauss-Newton proceeds iteratively by solving the linear equation:

$$\nabla^2_{\boldsymbol{\theta}} J \Delta\boldsymbol{\theta} = -\nabla_{\boldsymbol{\theta}} J. \tag{27}$$

However, in our case the Hessian, $\nabla^2_{\boldsymbol{\theta}} J$, is singular due to the ambiguity directions with rank $n - r$. Hence we take steps in the direction:

$$\Delta\boldsymbol{\theta} = -(\nabla^2_{\boldsymbol{\theta}} J)^-_{n-r} \nabla J, \tag{28}$$

which proceeds orthogonally towards the manifold $\mathcal{M}$. This is called free-gauge minimization. To implement LM we add a gradient term.

At the solution, $\boldsymbol{\theta} \in \mathcal{M}$, the covariance of the ML estimation of shape and motion parameters is obtained as:

$$V[\boldsymbol{\theta}] = E\{\Delta\boldsymbol{\theta}\Delta\boldsymbol{\theta}^\top\} = 2(\nabla^2_{\boldsymbol{\theta}} J)^-_{n-r} \tag{29}$$

$$= \frac{1}{2}(\nabla^{\mathcal{T}}_{\boldsymbol{\theta}} \mathbf{\Pi}(\boldsymbol{\theta}) V_p^{-1} \nabla^{\mathcal{T}}_{\boldsymbol{\theta}} \mathbf{\Pi}(\boldsymbol{\theta})^\top)^-_{n-r} \tag{30}$$

It can be shown that this is identical to the normal covariance expression in equation (14), $V[\boldsymbol{\theta}] \equiv V_{\perp\mathcal{M}}[\boldsymbol{\theta}]$, and not just up to a geometric equivalence, and so we use this as an alternate expression to for the normal covariance.

4.2 Efficient Covariance Estimation

The calculation of the generalized inverse in equations (28) and (30) involves use of SVD which takes $O(n^3)$ operations, and so for many feature points or images is slow. The Hessian often has sparse structure and when it is multiplied by the gradient, as in LM, the generalized inverse can be avoided and efficient minimization methods for J have been proposed [1, 2]. Here, however, we not only want a fast LM method, but also an efficient method to estimate the full covariance. We propose an efficient method in this section.

Let us assume that our parameter vector is divided into a shape and a motion part, $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_m$ respectively, such that $\boldsymbol{\theta} = (\boldsymbol{\theta}_s^\top, \boldsymbol{\theta}_m^\top)^\top$. The Hessian is then split into its shape and motion components:

$$\nabla^2_{\boldsymbol{\theta}} J = \begin{pmatrix} \nabla^2_{\boldsymbol{\theta}_s} J & \nabla_{\boldsymbol{\theta}_{sm}} J \\ \nabla_{\boldsymbol{\theta}_{ms}} J & \nabla^2_{\boldsymbol{\theta}_m} J \end{pmatrix} = \begin{pmatrix} U & W \\ W^\top & V \end{pmatrix}. \tag{31}$$

When noise in the feature points specified by V_p are independent of each other, U and V are full rank[3] and sparse with $O(N)$ and $O(M)$ non-zero elements respectively, where N is the number of features and M is the number of images [2].

[3] U is full rank for affine and perspective projection, but not when homogeneous coordinates are used as the general projective case, but then we do not obtain Euclidean shape.

The cross-term matrix W is not sparse however, and so applying a standard sparse techniques will not reduce the complexity of determining the generalized inverse.

First we define the full rank matrix T as follows:

$$T = \begin{pmatrix} I & 0 \\ -W^\top U^{-1} & I \end{pmatrix} \tag{32}$$

and obtain the block diagonal matrix:

$$T\nabla^2_{\boldsymbol{\theta}} J T^\top = \begin{pmatrix} U & 0 \\ 0 & V - W^\top U^{-1} W \end{pmatrix}. \tag{33}$$

Then we define the covariance $V_T[\boldsymbol{\theta}]$ by:

$$\begin{aligned} V_T[\boldsymbol{\theta}] &= T^\top (T\nabla^2_{\boldsymbol{\theta}} J T^\top)^-_{n-r} T \\ &= T^\top \begin{pmatrix} U^{-1} & 0 \\ 0 & (V - W^\top U^{-1} W)^-_{m-r} \end{pmatrix} T, \end{aligned} \tag{34}$$

where $m = 6M$ is the number of motion parameters. This can be obtained in $O(N^2M + M^3)$ operations which, when when the number of images is small (i.e. $M \ll N$), is much faster than the original SVD which is $O(N^3 + M^3)$.

In order for $V_T[\boldsymbol{\theta}]$ to be a valid description of the uncertainty, we must show that it is geometrically equivalent to $V_{\perp\mathcal{M}}[\boldsymbol{\theta}]$. Let $A = \frac{1}{2}\nabla^2_{\boldsymbol{\theta}} J$ be half the Hessian, and consider the equation:

$$A\mathbf{x} = \mathbf{u} \tag{35}$$

where $\mathbf{u}$ is in the column space of A. The general solution is a combination of a unique particular solution, $\mathbf{x}_p = A^-\mathbf{u}$, in the column space of A, and a homogeneous solution, $\mathbf{x}_h$, which is any vector in the nullspace of A, i.e. $A\mathbf{x}_h = 0$. We left multiply equation (35) by T and rearrange to obtain:

$$(TAT^\top)T^{-\top}\mathbf{x} = T\mathbf{u}. \tag{36}$$

Then changing variables: $\mathbf{y} = T^{-\top}\mathbf{x}$, and solving for $\mathbf{y}$ we obtain: $\mathbf{y} = (TAT^\top)^- T\mathbf{u} + \mathbf{y}_h$ where $(TAT^\top)\mathbf{y}_h = 0$. Now transforming back to $\mathbf{x}$ we can decompose the solution into the particular and homogeneous parts:

$$\mathbf{x} = T^\top (TAT^\top)^- T\mathbf{u} + T^\top \mathbf{y}_h = \mathbf{x}_p + \mathbf{x}_h, \tag{37}$$

where $\mathbf{x}_p = A^-\mathbf{u}$ is the particular solution obtained in equation (35). It is easy to see that $T^\top \mathbf{y}_h$ is in the nullspace of A, and hence $T^\top (TAT^\top)^- T\mathbf{u} = \mathbf{x}_p + \mathbf{x}'_h$ for some vector $\mathbf{x}'_h$ in the nullspace of A.

We now apply the geometric equivalence test to $V_{\perp\mathcal{M}}[\boldsymbol{\theta}] = A^-$ and $V_T[\boldsymbol{\theta}] = T^\top (TAT^\top)^- T$. The change of constraint Jacobian is the identity: $\frac{\partial \boldsymbol{\theta}_c}{\partial \boldsymbol{\theta}_{c'}} = I$, and the orthogonal component to the tangent space of $\mathcal{M}$, $T_{\boldsymbol{\theta}}[\mathcal{M}]^\perp$, is spanned by the columns of A and so $\mathbf{u}$ is any vector in the column space of A. Applying the equivalence relationship we obtain:

$$\mathbf{u}^\top (A^- - T^\top (TAT^\top)^- T)\mathbf{u} = \mathbf{u}^\top (\mathbf{x}_p - \mathbf{x}_p - \mathbf{x}'_h) = \mathbf{u}^\top (-\mathbf{x}'_h) = 0, \tag{38}$$

for all $\mathbf{u}$ in the column and row space of A, since $\mathbf{x}'_h$ is in the nullspace. We thus conclude that $V_T[\boldsymbol{\theta}]$ can be efficiently estimated and is geometrically equivalent to the normal covariance $V_{\perp\mathcal{M}}[\boldsymbol{\theta}]$.

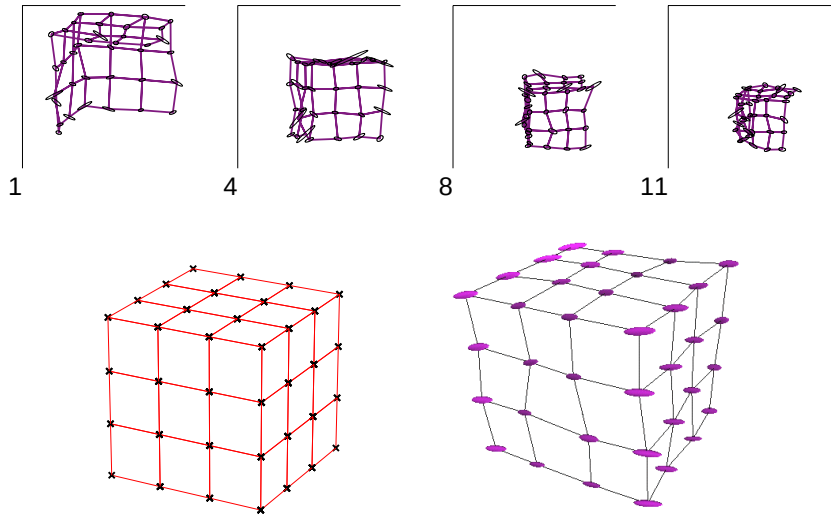

Fig. 3. Four images of an eleven image sequence with significant noise added and the scaled standard deviation of each point illustrated with an ellipse. (The lines connecting points are only present for viewing). The synthetic object is shown bottom left. The optimal reconstruction, given the noise estimates, is shown on the right with uncertainty ellipsoids. These ellipsoids, corresponding to the 3×3 block diagonal elements of a full shape covariance, are significantly correlated as shown in the full covariance matrix in Figure 5.

5 Results

We give some sample synthetic and real results illustrating our uncertainty modeling. A set of features in an image sequence with known correspondences is shown in Figure 3. The synthetic object is also shown along with a sample optimal reconstruction and ellipsoids illustrating feature-based uncertainty. The individual feature uncertainties are strongly correlated as illustrated in the subsequent Figures.

The normal covariance for this shape and motion recovery is shown in Figure 4. This contains a full description of the uncertainty in the features, but to experimentally confirm it using Monte Carlo simulation requires that we select a gauge such as that in equation (8). In Figure 5 we show the predicted covariance obtained by projecting the normal covariance into this gauge using equation (17). Even though the normal covariance and the predicted covariance

have significantly different values and correlations, they contain the same geometric uncertainty (as they are equivalent under the Geometric Equivalence Relationship) and will give the same predictions for uncertainties of gauge invariants. Figure 5 also contains the Monte Carlo covariance estimate in this gauge, involving 400 SFM reconstruction runs. It is very similar to the predicted covariance confirming that our uncertainty model is correct. An easier way to visually compare the covariances is to plot the square root of their diagonal elements. This gives the net standard deviation in each parameter in this gauge as illustrated in Figure 6.

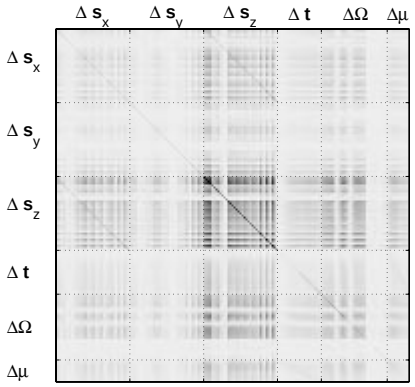

Fig. 4. The predicted normal covariance matrix giving us the geometric uncertainty of the reconstructed synthetic object. The scaled absolute value is shown by the darkness of the shading. Here weak perspective was used and μ is the recovered scale for each image. We note that it can be altered by adding components in the tangent plane to $\mathcal{M}$ without changing the underlying uncertainty, as we see in Figure 5.

The problem with the shape and motion covariance plots is their dependence on choice of gauge. Gauge invariants, however, will give us unambiguous measures for the uncertainty of the results. We chose two invariants on our synthetic object: an angle between two lines and the ratio of two lengths. Their statistics are shown in Table 1, confirming very good matching between predicted uncertainty and actual uncertainty.

Table 1. Predicted and measured values, along with their uncertainties in standard deviations, of two gauge independent properties of the synthetic object in Figure 3: (left) the angle between two lines, and (right) the ratio of two lengths.

Angle	*Mean*	*Uncertainty*	**Ratio**	*Mean*	*Uncertainty*
Predicted:	90.11°	±2.10°	Predicted:	0.9990	±0.0332
Recovered:	90.02°	±2.10°	Recovered:	1.0005	±0.0345

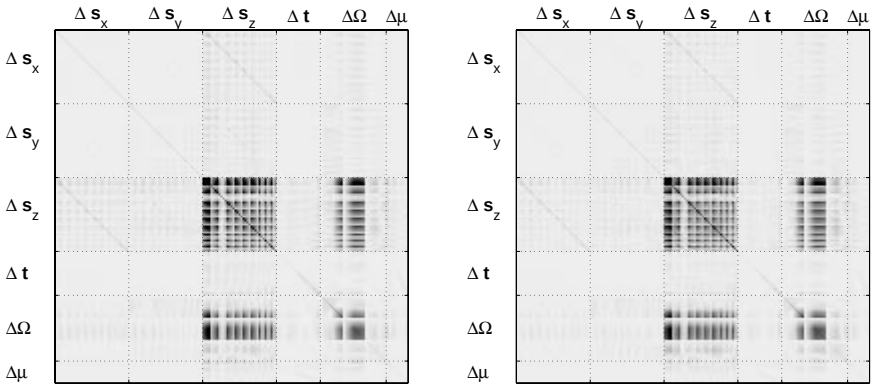

Fig. 5. (Left diagram) The predicted covariance in an arbitrary gauge, see equation (8). We note that the values and correlations are significantly different from the normal covariance in Figure 4, and yet it still contains the same geometric uncertainty. The Monte Carlo estimation of covariance in this gauge is shown on the right. It shows close similarity to the predicted covariance in this gauge as can also be seen in Figure 6.

Next we show results for a real image sequence of a chapel in Figure 7 along with the reconstructed shape from SFM. The feature correspondences were determined manually. Not only can we obtain a texture-mapped reconstruction, we can also obtain measurements of similarity invariant properties such as angles with their uncertainties. We found the angle and its uncertainty (in standard deviations) between two walls separated by a buttress: $117° \pm 3.2°$, as well as two other angles on the chapel: $46.2° \pm 2.1°$ and $93.2° \pm 2.6°$ as described in the Figure caption. These quantities are exact and not only up to an unknown transformation. We believe reporting this uncertainty measure is essential for most quantitative analyzes of the shape, and can only be done for gauge invariant properties.

6 Concluding Remarks

We have addressed the question of uncertainty representation when parameter indeterminacies exist in estimation problems. The shape and motion parameters estimated by SFM contain inherent indeterminacies. Hence to apply perturbation analysis, these parameters are first constrained by a gauge and the covariance is estimated in this gauge. Unfortunately the choice of gauge will have significant effects on the uncertainties of the parameters, as illustrated in our results. These effects, however, are "non-physical" and do not correspond to changes in the actual geometric uncertainty which is unaffected by an arbitrary choice of gauge. Thus shape and motion parameter uncertainties contain artifacts of the choice of gauge. The uncertainties of gauge invariant parameters, however, are not affected by these indeterminacies and hence correspond directly

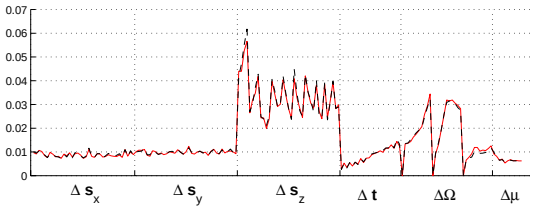

Fig. 6. The square root of the diagonal elements of the covariances in Figure 5 are shown here. This gives the net standard deviation of each parameter in the experimental gauge (8) obtained from the diagonal of the covariance. The solid line is the experimentally measured value and the dashed line is our prediction from the projected normal covariance.

to the inherent geometric uncertainty. They thus provide unambiguous measures for the solution uncertainty.

We derived a Geometric Uncertainty Relationship which permits us to compare the geometric uncertainty contained in covariances described under different parametrizations and gauges. Using this relationship we showed that the normal covariance, whose estimation does not need explicit gauge constraints, fully describes the solution uncertainty. We were also able to derive an efficient estimation method for the solution covariance. Using the Geometric Uncertainty Relationship, we showed that this estimate also fully captures the solution uncertainty. Gauge invariant uncertainties can be calculated by transforming this covariance.

References

1. K. B. Atkinson, *Close Range Photogrammetry and Machine Vision*, Whittles Publ., Caithness, Scotland, (1996).
2. R. Hartley, Euclidean reconstruction from uncalibrated views, *Proc. DARPA–ESPRIT Workshop App. Invariants in Comp. Vis.*, Portugal, 1993 187–202.
3. A. Heyden & K. Astrom, Euclidean reconstruction from image sequences with varying and unknown focal length and principal point, *Proc. Comp. Vision Patt. Recog.*, Peurto Rico, 1997 438–443.
4. T. Kanade & D. D. Morris, Factorization methods for structure from motion, *Phil. Trans. R. Soc. Long. A*, 1998 1153–1173.
5. K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier, Amsterdam, (1996).
6. K. Kanatani, Geometric information criterion for model selection, *Int. J. Comp. Vision* (1998), **26**(3), 171–189.
7. K. Kanatani & D. D. Morris, Gauges and gauge transformations in 3-D reconstruction from a sequence of images, *To appear in ACCV 2000*.
8. L. L. Kontsevich, M. L. Kontsevich & A. K. Shen, Two algorithms for reconstructing shapes, *Avtometriya* (1987), (5), 76–81, (Transl. Optoelec., Instrum. Data Proc., no. 5. 76–81, 1987).

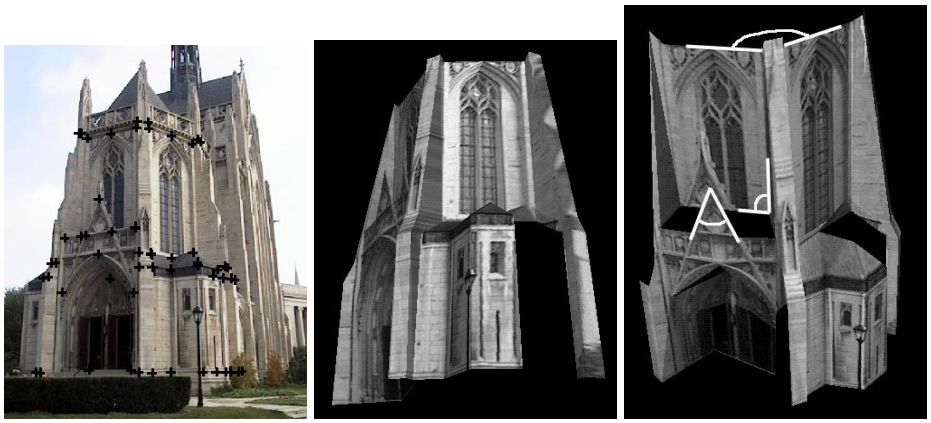

Fig. 7. An image from a 6 image sequence of a chapel is shown on the left with features registered by hand. The reconstruction is on the right. We can obtain quantitative measures and uncertainties from this reconstruction. In this case we estimated the angle between to walls separated by a buttress, and two other angles as illustrated on the far right. The values (anti-clockwise from the top) are: $117^\circ \pm 3.2^\circ$, $46.2^\circ \pm 2.1^\circ$ and $93.2^\circ \pm 2.6^\circ$.

9. D. D. Morris & T. Kanade, A unified factorization algorithm for points, line segments and planes with uncertainty models, *Proc. Sixth Int. Conf. Comp. Vision*, Bombay, India, 1998 696–702.
10. C. Poelman & T. Kanade, A paraperspective factorization method for shape and motion recovery, *IEEE Trans. Patt. Anal. Mach. Intell.* (Mar. 1997), **19**(3), 206–18.
11. L. Quan & T. Kanade, Affine structure from line correspondances with uncalibrated affine cameras, *IEEE Trans. Patt. Anal. Mach. Intell.* (1997), **19**(8), 834–45.
12. P. Sturm & B. Triggs, A factorization based algorithm for multi-image projective structure and motion, *Proc. Euro. Conf. Comp. Vision*, Cambridge, UK, 1996 709–20.
13. R. Szeliski & S. B. Kang, Recovering 3D shape and motion from image streams using non-linear least squares, *J. of Visual Comm. and Image Rep.* (Mar. 1994), **5**(1), 10–28.
14. R. Szeliski & S. B. Kang, Shape ambiguities in structure from motion, *IEEE Trans. Patt. Anal. Mach. Intell.* (May 1997), **19**(5), 506–512.
15. J. I. Thomas, A. Hanson & J. Oliensis, Refining 3D reconstructions: A theoretical and experimental study of the effect of cross-correlations, *CVGIP* (Nov. 1994), **60**(3), 359–370.
16. J. Weng, N. Ahuja & T. Huang, Optimal motion and structure estimation, *IEEE Trans. Patt. Anal. Mach. Intell.* (1993), **15**(9), 864–884.