Gauges and Gauge Transformations for Uncertainty Description of Geometric Structure with Indeterminacy

Kenichi Kanatani, *Senior Member, IEEE,* and Daniel D. Morris

Abstract—This paper presents a consistent theory for describing indeterminacy and uncertainty of three–dimensional (3-D) reconstruction from a sequence of images. First, we give a group-theoretical analysis of *gauges* and *gauge transformations*. We then discuss how to evaluate the reliability of the solution that has indeterminacy and extend the Cramer–Rao lower bound to incorporate internal indeterminacy. We also introduce the *free-gauge approach* and define the *normal form* of a covariance matrix that is independent of particular gauges. Finally, we show simulated and real-image examples to illustrate the effect of gauge freedom on uncertainty description.

Index Terms—Computer vision, Cramer–Rao lower bound, gauge transformation, geometric indeterminacy, Lie group theory, statistical estimation, uncertainty description.

I. INTRODUCTION

RECONSTRUCTING the three-dimensional (3-D) shape of a scene from a sequence of images is one of the central themes of *computer vision* [1], [2], [5], [6]. It was originally motivated toward industrial applications such as automatic manufacturing and inspection by robots, autonomous robots and land vehicles, and military uses such as autonomous surveillance and reconnaissance. Recently, interest in media and Internet applications has been quickly growing: visually presenting 3-D shapes of real objects through Internet, generating virtual images by embedding graphics objects in real scenes or real objects in graphics scenes, and machine understanding of the video content for intelligent communications, to name a few.

There is, however, another important application of computer vision: accurately measuring the 3-D shape using camera images for objects for which measurement is otherwise difficult, typical examples being 3-D measurement of urban architectural scenes and ancient archaeological structures [15]. One hurdle in such tasks is the need to address the accuracy quantitatively,

but this has not attracted much attention in the past. Typically, the accuracy has been evaluated empirically by simulated and real-image experiments or analytically by first-order error propagation [23]. This, however, hides an important fact: *3-D reconstruction from images involves indeterminacy.*

There are two sources of indeterminacy in parameterized 3-D shape reconstruction. One is the inherent physical indeterminacy caused by loss of information during projection onto the camera; e.g., scale is indeterminate for a perspective camera. The second source is the overparameterization of the problem: more than the minimal number of parameters are included for simplicity as well as symmetry of representation; e.g., a shape model with a vector representing each point specifies the absolute orientation and translation of the object though these may be indeterminate.

These indeterminacies can be removed by imposing normalization constraints; e.g., we can fix the coordinate origin to a particular point of the object and normalize the size of something to unit length. But then the fixed or normalized parameters have no uncertainty by definition, and the uncertainty of all other parameters is altered.

This fact makes it difficult to evaluate the performance of a 3-D reconstruction algorithm or compare it with another algorithm. For example, if simulation results are in good agreement with the true structure, one is tempted to claim that the method is accurate. This agreement, however, may be merely due to the normalization involved. For instance, the computed object shape and the true shape may agree well if the object size is normalized to unity, but they may not agree well if the camera displacement is normalized to unity. This kind of arbitrary or unfair evaluation of 3-D reconstruction algorithms has been found in many demonstrations.

The basic message of this paper is that we cannot speak of the uncertainty of a particular parameter in absolute terms unless it is *invariant* to normalization conditions. While uncertainty modeling has been a recent subject of research [13], [19], [23], the effects of normalization conditions have mostly been ignored.

In this paper, we call transformations of parameters caused by changing normalization conditions *gauge transformations* and particular choices of normalization *gauges*; these terms are borrowed from physics, but their meaning is not directly related to those in physics. Only very recently have these concepts come to be recognized as playing an important role in 3-D reconstruction from images [8], [10], [11], [13].

K. Kanatani was with the Department of Computer Science, Gunma University, Gunma, Japan. He is now with the Department of Information Technology, Okayama University, Okayama 700-8530, Japan (e-mail: kanatani@suri.it.okayama-u.ac.jp).

D. D. Morris is with the Robotics Institute, Carnegie-Mellon University, Pittsburgh, PA 1523-3890 USA (e-mail: ddmorris@ri.cmu.edu).

In the following, we present a consistent theory for describing indeterminacy and uncertainty of 3-D reconstruction from images. First, we give a formal description of the 3-D reconstruction problem by combining Lie group theory and the statistical estimation theory. We then extend the Cramer–Rao lower bound to incorporate internal indeterminacy. Finally, we introduce the *free-gauge approach* and define the *normal form* of a covariance matrix that is independent of particular gauges. We show simulated and real-image examples to illustrate the effect of gauge freedom on uncertainty description.

Our analytical tool is the local linear approximation, a standard technique in differential geometry: we introduce local coordinates and project perturbations of quantities onto tangent spaces, assuming that the noise is small.

II. PARAMETRIC RECONSTRUCTION OF 3-D SHAPE

We take *3-D reconstruction* to mean the computation of the parameters that describe the 3-D shape and motion of an object from its camera images. In this paper, we focus on the *feature-based* approach: we track the motion of easily identifiable feature points, such as corners and markers, over a sequence of images and compute their 3-D positions using the knowledge of the camera projection model [15], [20].

Suppose we track N rigidly moving feature points over M images. Let $(x_{\kappa\alpha}, y_{\kappa\alpha})$ be the image coordinates of the αth point in the κth frame. Here, we adopt the *camera-centered* description, assuming that an object is moving in the scene relative to a stationary camera, but the subsequent analysis is essentially the same if we view the camera as moving and taking images of a stationary scene.

We identify the camera coordinate system with a global reference frame and fix an arbitrary object coordinate system to the object. Let $\boldsymbol{t}_\kappa$ and $\{\boldsymbol{i}_\kappa, \boldsymbol{j}_\kappa, \boldsymbol{k}_\kappa\}$ be, respectively, its origin and orthonormal basis vectors in the κth frame. Let $(s_{\alpha 1}, s_{\alpha 2}, s_{\alpha 3})$ be the coordinates of the αth feature point with respect to the object coordinate system. We define

$$\boldsymbol{R}_\kappa = (\boldsymbol{i}_\kappa \quad \boldsymbol{j}_\kappa \quad \boldsymbol{k}_\kappa) \qquad \boldsymbol{s}_\alpha = \begin{pmatrix} s_{\alpha 1} \\ s_{\alpha 2} \\ s_{\alpha 3} \end{pmatrix}. \tag{1}$$

Since $\{\boldsymbol{i}_\kappa, \boldsymbol{j}_\kappa, \boldsymbol{k}_\kappa\}$ are a right-handed orthonormal bases, $\boldsymbol{R}_\kappa$ is a rotation matrix. We call $\{\boldsymbol{t}_\kappa, \boldsymbol{R}_\kappa\}$ the *motion parameters* and $\{\boldsymbol{s}_\alpha\}$ the *shape vectors*.

III. CAMERA IMAGING GEOMETRY

The 3-D positions of the tracked feature points are reconstructed by inverting the camera imaging equations, which we describe in this section.

Suppose a point (X, Y, Z) with respect to the camera coordinate system is projected to a point (x, y) on the image plane. We write this mapping as $\Pi: \mathcal{R}^3 \to \mathcal{R}^2$ ($\mathcal{R}$ denotes the set of real numbers) and call it the *camera model*. The mapping Π may be different from frame to frame. A point with object coordinates $(s_{\alpha 1}, s_{\alpha 2}, s_{\alpha 3})$ is a point $\boldsymbol{R}_\kappa \boldsymbol{s}_\alpha + \boldsymbol{t}_\kappa$ with respect to the camera

coordinate system in the κth frame. Hence, the image coordinates $(x_{\alpha\kappa}, y_{\alpha\kappa})$ are expressed as a function of $\boldsymbol{t}_\kappa$, $\boldsymbol{R}_\kappa$, and $\boldsymbol{s}_\alpha$ in the form

$$\begin{pmatrix} x_{\alpha\kappa} \\ y_{\alpha\kappa} \end{pmatrix} = \Pi_\kappa[\boldsymbol{R}_\kappa \boldsymbol{s}_\alpha + \boldsymbol{t}_\kappa] \tag{2}$$

where Π_κ is the camera model for the κth frame. We call the mapping $\mathcal{P}_\kappa: SO(3) \times \mathcal{R}^3 \times \mathcal{R}^3 \to \mathcal{R}^2$ the *projection model* for the κth frame, where $SO(3)$ denotes the group of 3-D rotations.

Example 1: For an *orthographic camera*, we identify the xy plane with the image plane and the z-axis with the optical axis of the lens. A point (X, Y, Z) in the scene is projected onto an image point (x, y) such that $x = X$ and $y = Y$. The projection model is given by

$$\begin{pmatrix} x_{\alpha\kappa} \\ y_{\alpha\kappa} \end{pmatrix} = \Pi(\boldsymbol{R}_\kappa \boldsymbol{s}_\alpha + \boldsymbol{t}_\kappa) \tag{3}$$

where $\Pi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. If we incorporate the effect of scale change due to translations in the depth orientation in the form

$$\begin{pmatrix} x_{\alpha\kappa} \\ y_{\alpha\kappa} \end{pmatrix} = \mu_\kappa \Pi(\boldsymbol{R}_\kappa \boldsymbol{s}_\alpha + \boldsymbol{t}_\kappa) \tag{4}$$

we obtain the *weak perspective* (or *scaled-orthographic*) camera model [14].

Example 2: For a *perspective* (or *pinhole*) *camera*, we identify the origin O with the center of the lens and the z-axis with its optical axis. A point (X, Y, Z) in the scene is projected onto an image point (x, y) such that $x = fX/Z$ and $y = fY/Z$, where f is a constant called the *focal length*. The projection model is given by

$$\begin{pmatrix} x_{\alpha\kappa} \\ y_{\alpha\kappa} \end{pmatrix} = f_\kappa \Pi[\boldsymbol{R}_\kappa \boldsymbol{s}_\alpha + \boldsymbol{t}_\kappa] \tag{5}$$

where f_κ is the focal length for the κth frame. The operation Π is defined by

$$\Pi\left[\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}\right] = \begin{pmatrix} X/Z \\ Y/Z \end{pmatrix}$$

(not defined for $Z = 0$). The projection model of a more general *projective camera* is

$$\begin{pmatrix} x_{\alpha\kappa} \\ y_{\alpha\kappa} \end{pmatrix} = \Pi[\boldsymbol{K}_\kappa(\boldsymbol{R}_\kappa \boldsymbol{s}_\alpha + \boldsymbol{t}_\kappa)] \tag{6}$$

where $\boldsymbol{K}_\kappa$ is an upper triangular matrix called the *camera matrix* of the κth frame. We can parameterize the camera matrix $\boldsymbol{K}$ in the form

$$\boldsymbol{K} = \begin{pmatrix} f & -f\cot\theta & u_0 \\ 0 & f/\alpha\sin\theta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7}$$

where (u_0, v_0) is the *principal point* (the point that corresponds to the optical axis orientation in the image). The constants $f, \theta,$ and α are called the *focal length*, the *skew angle*, and the *aspect ratio*, respectively [2].

IV. PARAMETER SPACE

Now that the camera model for a single point in a single image has been defined, we extend it to all features in all images. We

stack the two-dimensional (2-D) vector (2) for all the M frames vertically and define a $2M$-dimensional vector $\boldsymbol{p}_\alpha$ so that the image motion of the αth point can be identified with a single point $\boldsymbol{p}_\alpha$ in a $2M$-dimensional space $\mathcal{R}^{2M}$, which we call the *data space*.

Let $\boldsymbol{\theta}$ be a vector that stacks the shape vectors $\{\boldsymbol{s}_\alpha\}$, the motion parameters $\{\boldsymbol{t}_\kappa, \boldsymbol{R}_\kappa\}$, and the parameters that specify the camera models $\{\Pi_\kappa\}$ for all the frames vertically. We denote the domain of $\boldsymbol{\theta}$ by $\mathcal{M}$ and call it the *parameter space*, which is a manifold in general. The image motion $\boldsymbol{p}_\alpha$ is a function of the parameter vector $\boldsymbol{\theta}$

$$\boldsymbol{p}_\alpha = \mathcal{P}_\alpha[\boldsymbol{\theta}] \tag{8}$$

which defines the projection model of the "$2M$-dimensional camera."

Example 3: For the orthographic camera model, the vector $\boldsymbol{\theta}$ has dimension $6M + 3N$ (the rotation matrix $\boldsymbol{R}_\kappa$ and the translation vector $\boldsymbol{t}_\kappa$ both have three degrees of freedom). Topologically, the parameter space $\mathcal{M}$ is homeomorphic to the direct product[1] of $\mathcal{R}^{3M+3N}$ and M copies of $SO(3)$

$$\mathcal{M} \cong \overbrace{SO(3) \times \cdots \times SO(3)}^{M} \times \mathcal{R}^{3M+3N}. \tag{9}$$

V. STATISTICAL NOISE MODEL

In real applications, the feature points can be determined only to a finite accuracy due to image digitization. Moreover, they are corrupted by unknown elements due to unmodeled aspects of the imaging process and the image processing operations for feature detection. We model this uncertainty as follows.

Let $\bar{\boldsymbol{p}}_\alpha$ be the true value of $\boldsymbol{p}_\alpha$, and write $\boldsymbol{p}_\alpha = \bar{\boldsymbol{p}}_\alpha + \Delta\boldsymbol{p}_\alpha$. We assume that the noise term $\Delta\boldsymbol{p}_\alpha$ is a Gaussian random variable; it may not be independent for different images, and its distribution may be different from point to point. We define their covariance matrices by

$$V[\boldsymbol{p}_\alpha, \boldsymbol{p}_\beta] = E\left[\Delta\boldsymbol{p}_\alpha \Delta\boldsymbol{p}_\beta^\top\right] \tag{10}$$

where $E[\cdot]$ denotes expectation.

In practical applications, the absolute magnitude of image noise is very difficult to predict *a priori*, but its geometric characteristics such as homogeneity/inhomogeneity and isotropy/anisotropy can be predicted relatively easily. For example, if we use template matching for finding corresponding points, the uncertainty of matching is characterized by the variation in the matching residual around the detected point [3], [12], [17], [18]. Hence, we assume that $V[\boldsymbol{p}_\alpha, \boldsymbol{p}_\beta]$ is known *up to scale* and decompose it into the known *normalized covariance matrix* $V_0[\boldsymbol{p}_\alpha, \boldsymbol{p}_\beta]$ and an unknown *noise level* ϵ as follows [6]:

$$V[\boldsymbol{p}_\alpha, \boldsymbol{p}_\beta] = \epsilon^2 V_0[\boldsymbol{p}_\alpha, \boldsymbol{p}_\beta]. \tag{11}$$

This decomposition can be defined arbitrarily but is fixed once for all.

---

[1]If algebraic operations are defined among the elements of $\mathcal{M}$ so that $\mathcal{M}$ becomes a Lie group, the symbol $\times$ is often called the *semidirect product*. Here, however, we view $\mathcal{M}$ simply as a manifold and do not define operations among its elements.

Define $2M \times 2M$ matrices $\boldsymbol{W}_{\alpha\beta}, \alpha, \beta = 1, \ldots, N$, by

$$\begin{pmatrix} \boldsymbol{W}_{11} & \cdots & \boldsymbol{W}_{1N} \\ \vdots & & \vdots \\ \boldsymbol{W}_{N1} & \cdots & \boldsymbol{W}_{NN} \end{pmatrix}$$
$$= \begin{pmatrix} V_0[\boldsymbol{p}_1, \boldsymbol{p}_1] & \cdots & V_0[\boldsymbol{p}_1, \boldsymbol{p}_N] \\ \vdots & & \vdots \\ V_0[\boldsymbol{p}_N, \boldsymbol{p}_1] & \cdots & V_0[\boldsymbol{p}_N, \boldsymbol{p}_N] \end{pmatrix}^{-1}. \tag{12}$$

An optimal solution is obtained by *maximum-likelihood estimation*, often referred to as *bundle adjustment* [22]: we minimize

$$J(\boldsymbol{\theta}) = \sum_{\alpha, \beta=1}^{N} (\boldsymbol{p}_\alpha - \mathcal{P}_\alpha[\boldsymbol{\theta}], \boldsymbol{W}_{\alpha\beta}(\boldsymbol{p}_\beta - \mathcal{P}_\beta[\boldsymbol{\theta}])) \tag{13}$$

where and throughout this paper we denote the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ by $(\boldsymbol{a}, \boldsymbol{b})$. The global minimum of this function provides us with our optimal shape and motion estimates.

VI. GAUGE TRANSFORMATIONS

In general, the solution that minimizes $J(\boldsymbol{\theta})$ is not unique. Rather, there is a continuum of shape and motion parameters that all achieve the same global minimum. There are two sources of this indeterminacy:

Frame indifference: Absolute translations and orientations of objects cannot be determined uniquely.

Projection insensitivity: Different 3-D configurations can be projected to the same image.

If there exists a transformation g of the parameter space $\mathcal{M}$ such that

$$J(\boldsymbol{\theta}) = J(g\boldsymbol{\theta}), \qquad \forall \boldsymbol{\theta} \in \mathcal{M} \tag{14}$$

we call g a *gauge transformation*. The set of all such transformations forms a group $\mathcal{G}$, which we call the *group of gauge transformations* or the *gauge group* for short.

Example 4: If the object coordinate system is rotated by $\boldsymbol{R}$ around its origin and then translated by $\boldsymbol{t} = (t_1, t_2, t_3)^\top$, where $\boldsymbol{t}$ and $\boldsymbol{R}$ are defined with respect to the original object coordinate system, the shape vectors $\{\boldsymbol{s}_\alpha\}$ and the motion parameters $\{\boldsymbol{t}_\kappa, \boldsymbol{R}_\kappa\}$ are transformed in the form

$$\boldsymbol{s}'_\alpha = \boldsymbol{R}^\top(\boldsymbol{s}_\alpha - \boldsymbol{t}) \qquad \boldsymbol{t}'_\kappa = \boldsymbol{R}_\kappa\boldsymbol{t} + \boldsymbol{t}_\kappa \qquad \boldsymbol{R}'_\kappa = \boldsymbol{R}_\kappa\boldsymbol{R}. \tag{15}$$

Example 5: For the orthographic camera model, we see from (3) that if $\{\boldsymbol{s}_\alpha\}$ and $\{\boldsymbol{t}_\kappa, \boldsymbol{R}_\kappa\}$ are a solution, so are $\{\boldsymbol{s}_\alpha\}$ and $\{\boldsymbol{t}_\kappa + d_\kappa\boldsymbol{k}, \boldsymbol{R}_\kappa\}$ for all $d_\kappa \in \mathcal{R}$, where and throughout this paper we let $\boldsymbol{k} = (0, 0, 1)^\top$. We also see from (3) that if $\{\boldsymbol{s}_\alpha\}$ and $\{\boldsymbol{t}_\kappa, \boldsymbol{R}_\kappa\}$ are a solution, so are $\{-\boldsymbol{s}_\alpha\}$ and $\{-\boldsymbol{J}\boldsymbol{t}_\kappa, \boldsymbol{J}\boldsymbol{R}_\kappa\}$, where $\boldsymbol{J} = \mathrm{diag}(-1, -1, 1)$ ($180°$ rotation around the Z-axis). The two solutions are *mirror images* of each other; this ambiguity is known as the *Necker cube phenomenon*. For the perspective camera model, we see from (5) that if $\{\boldsymbol{s}_\alpha\}$ and $\{\boldsymbol{t}_\kappa, \boldsymbol{R}_\kappa\}$ are a solution, so are $\{s\boldsymbol{s}_\alpha\}$ and $\{s\boldsymbol{t}_\kappa, \boldsymbol{R}_\kappa\}$ for all scale $s \neq 0$. Negative scaling produces mirror images.

VII. INFINITESIMAL GAUGE GENERATORS

We now provide a local description of our gauge transformation. An element $\gamma \in \mathcal{G}$ close to the identity element of $\mathcal{G}$ is

called an *infinitesimal gauge transformation*. It maps $\boldsymbol{\theta}$ in the form

$$\gamma\boldsymbol{\theta} = \boldsymbol{\theta} + \boldsymbol{D}(\boldsymbol{\theta}) + \cdots \qquad (16)$$

where $\boldsymbol{D}(\cdot)$ is a vector operator called the *infinitesimal gauge generator* of γ. The set of such generators is a linear space,[2] which we denote by $\mathcal{D}_\theta(\mathcal{M})$. It can be proved that for every element $\boldsymbol{D}(\cdot) \in \mathcal{D}_\theta(\mathcal{M})$ there exists a gauge transformation γ that has $\boldsymbol{D}(\cdot)$ as its generator [4]. If the linear space $\mathcal{D}_\theta(\mathcal{M})$ of gauge generators is r-dimensional, $\mathcal{G}$ is an r-dimensional Lie group. We call the dimension r the *degree of gauge freedom*.

Example 6: The rotation around a unit vector $\boldsymbol{l}$ by a small angle $\Delta\Omega$ is written to a first approximation in the form $\boldsymbol{I} + \Delta\boldsymbol{\Omega} \times \boldsymbol{I}$, where $\boldsymbol{I}$ denotes the unit matrix and we let $\Delta\boldsymbol{\Omega} = \Delta\Omega\boldsymbol{l}$. Throughout this paper, the product $\boldsymbol{a} \times \boldsymbol{A}$ of a 3-D vector $\boldsymbol{a}$ and a 3×3 matrix is defined to be the matrix whose columns are the vector products of $\boldsymbol{a}$ with each of the columns of $\boldsymbol{A}$. We also define $\boldsymbol{A} \times \boldsymbol{a}$ to be $\boldsymbol{A}(\boldsymbol{a} \times \boldsymbol{I})^\top$, the matrix whose rows are the vector products of $\boldsymbol{a}$ with each of the rows of $\boldsymbol{A}$. For the orthographic camera model, the following generators define an infinitesimal transformation:

$$\boldsymbol{D}(\boldsymbol{s}_\alpha) = -\Delta\boldsymbol{\Omega} \times \boldsymbol{s}_\alpha - \Delta\boldsymbol{t}$$

$$\boldsymbol{D}(\boldsymbol{t}_\kappa) = \boldsymbol{R}_\kappa \Delta\boldsymbol{t} + \Delta d_\kappa \boldsymbol{k} \qquad \boldsymbol{D}(\boldsymbol{R}_\kappa) = -\boldsymbol{R}_\kappa \times \Delta\boldsymbol{\Omega}. \quad (17)$$

The operator $\boldsymbol{D}(\cdot)$ is linearly parameterized by $\Delta\boldsymbol{\Omega}$, $\Delta\boldsymbol{t}$, and $\{\Delta d_\kappa\}$, implying that the linear space $\mathcal{D}_\theta(\mathcal{M})$ is $(M+6)$-dimensional. For the perspective camera model, the following generators define an infinitesimal transformation:

$$\boldsymbol{D}(\boldsymbol{s}_\alpha) = -\Delta\boldsymbol{\Omega} \times \boldsymbol{s}_\alpha - \Delta\boldsymbol{t} + \Delta s \boldsymbol{s}_\alpha$$

$$\boldsymbol{D}(\boldsymbol{R}_\kappa) = -\boldsymbol{R}_\kappa \times \Delta\boldsymbol{\Omega} \qquad \boldsymbol{D}(\boldsymbol{t}_\kappa) = \boldsymbol{R}_\kappa \Delta\boldsymbol{t} + \Delta s \boldsymbol{t}_\kappa. \quad (18)$$

The linear space $\mathcal{D}_\theta(\mathcal{M})$ is seven-dimensional (7-D); it is parameterized by $\Delta\boldsymbol{\Omega}$, $\Delta\boldsymbol{t}$, and Δs.

VIII. GAUGES

The existence of gauge freedom implies that there exists a smooth manifold in the parameter space $\mathcal{M}$ such that any point $\boldsymbol{\theta}$ of it gives a solution for the same data. In order to disambiguate the solution, we define another manifold that intersects the manifold at a single point. The formal description is as follows.

We say that two values $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ are *geometrically equivalent* and write $\boldsymbol{\theta} \sim \boldsymbol{\theta}'$ if there exists an element $g \in \mathcal{G}$ such that $\boldsymbol{\theta}' = g\boldsymbol{\theta}$. This implies that the true parameter space is not $\mathcal{M}$ but the *quotient space* $\mathcal{M}/\mathcal{G}$ of $\mathcal{M}$ with respect to this equivalence relation. Let $\mathcal{M}_\theta$ be an element of $\mathcal{M}/\mathcal{G}$, i.e., a subset of $\mathcal{M}$ consisting of all points geometrically equivalent to $\boldsymbol{\theta}$

$$\mathcal{M}_\theta = \{\boldsymbol{\theta}' \mid \boldsymbol{\theta}' \sim \boldsymbol{\theta}\} \, (=\{g\boldsymbol{\theta} \mid g \in \mathcal{G}\}). \qquad (19)$$

If the degree of gauge freedom is r, $\mathcal{M}_\theta$ is an r-dimensional submanifold of $\mathcal{M}$ called the *leaf* associated with $\boldsymbol{\theta}$. The parameter space $\mathcal{M}$ is filled with leaves. Such a space is called a *foliation* or a *foliated manifold*.

A natural way to choose a unique value of $\boldsymbol{\theta}$ for each leaf is to assign r equations

$$c_1(\boldsymbol{\theta}) = 0, \, \ldots, \, c_r(\boldsymbol{\theta}) = 0. \qquad (20)$$

Each equation removes one degree of gauge freedom. We call these equations a *gauge condition*, or a *gauge* for short, if the following conditions are satisfied.

1) They are algebraically independent, defining a submanifold $\mathcal{C}$ of *codimension* r in $\mathcal{M}$; we call $\mathcal{C}$ the *gauge manifold*.

2) The gauge manifold $\mathcal{C}$ intersects *all* leaves $\mathcal{M}_\theta$ *transversally*[3] with each *connected component* at a single point.

3) For any $\boldsymbol{\theta} \in \mathcal{M}_\theta$ and $\boldsymbol{\theta}_\mathcal{C} = \mathcal{C} \cap \mathcal{M}_\theta$, there exists a *unique* element $g \in \mathcal{G}$ such that $\boldsymbol{\theta}_\mathcal{C} = g\boldsymbol{\theta}$.

Hereafter, we use the terms "gauge" and "gauge manifold" interchangeably, denoting both by $\mathcal{C}$. Introducing a gauge arbitrarily, we can find a solution $\boldsymbol{\theta}$ that minimizes (13) uniquely for each connected component of the leaf $\mathcal{M}_\theta$, but we cannot distinguish two solutions that belong to disjoint components.[4]

Example 7: For the orthographic camera model, we can choose the origin of the object coordinate system at the centroid of the feature points, align the axis orientation to the first camera frame, and let the Z-components of the translation $\boldsymbol{t}_\kappa$ be all zeros

$$\sum_{\alpha=1}^N \boldsymbol{s}_\alpha = \boldsymbol{0} \qquad \boldsymbol{R}_1 = \boldsymbol{I} \qquad (\boldsymbol{k}, \boldsymbol{t}_\kappa) = 0. \qquad (21)$$

We call this the *standard gauge*. This can be extended to the weak perspective camera model by adding

$$\sum_{\alpha=1}^N \|\boldsymbol{s}_\alpha\|^2 = 1. \qquad (22)$$

For the perspective camera model, we define the standard gauge as follows:

$$\boldsymbol{R}_1 = \boldsymbol{I} \qquad \boldsymbol{t}_1 = \boldsymbol{0} \qquad \|\boldsymbol{t}_2\| = 1. \qquad (23)$$

Example 8: If we differentiate (13) with respect to $\boldsymbol{t}_\kappa$, let the result be zero, and solve the resulting equation, we can express $\boldsymbol{t}_\kappa$ in terms of the remaining parameters. Substituting it into $J(\boldsymbol{\theta})$, we obtain a function of the remaining parameters; we call this description the *reduced model*. Similarly, we can eliminate $\{\boldsymbol{s}_\alpha\}$ and express $J(\boldsymbol{\theta})$ in terms of the motion parameters $\{\boldsymbol{t}_\kappa, \boldsymbol{R}_\kappa\}$ alone; we call this description the *epipolar model*.

IX. ESTIMATORS

Our goal is to construct a function $\hat{\boldsymbol{\theta}}(\{\boldsymbol{p}_\alpha\})$ of the data $\{\boldsymbol{p}_\alpha\}$ that approximates $\boldsymbol{\theta}$ as closely as possible. Such a mapping $\hat{\boldsymbol{\theta}}: \mathcal{R}^{2MN} \to \mathcal{M}$ is called an *estimator* of $\boldsymbol{\theta}$. The solution $\hat{\boldsymbol{\theta}}$ that minimizes (13) is called the *maximum-likelihood estimator*. However, we cannot speak of the "true solution" $\boldsymbol{\theta}$; we can only speak of the "true leaf" $\mathcal{M}_\theta$, since any two points on it are geometrically equivalent. We can make the solution unique by imposing a gauge $\mathcal{C}$; let us denote the resulting solution by $\boldsymbol{\theta}_\mathcal{C}$. An

---

[2]By introducing a product called *commutator*, we can make this linear space an algebra, which is isomorphic to the *Lie algebra* $\mathcal{L}(\mathcal{G})$ of $\mathcal{G}$.

[3]Intuitively, two manifolds intersect transversally if one is not tangent to the other [16].

[4]This occurs, for example, for the two mirror image solutions described in Example 5.

estimator that satisfies a gauge $\mathcal{C}$ is denoted by $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$. We assume that if noise does not exist, the estimator $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ coincides with the value $\boldsymbol{\theta}_{\mathcal{C}} \in \mathcal{M}_{\theta} \cap \mathcal{C}$ that satisfies the gauge $\mathcal{C}$ and belongs to the true leaf $\mathcal{M}_{\theta}$.

Let $T_{\theta_{\mathcal{C}}}(\mathcal{M})$ be the *tangent space* to $\mathcal{M}$ at $\boldsymbol{\theta}_{\mathcal{C}}$. This is an n-dimensional linear space, where and in the remainder of this paper n is the dimension of the parameter vector $\boldsymbol{\theta}$ and r is the degree of gauge freedom, i.e., the dimension of the linear space $\mathcal{D}_{\theta}(\mathcal{M})$ of gauge generators. Let $T_{\theta_{\mathcal{C}}}(\mathcal{C})$ be the tangent space to the gauge manifold $\mathcal{C}$ at $\boldsymbol{\theta}_{\mathcal{C}}$; it is an $(n-r)$-dimensional subspace of $T_{\theta_{\mathcal{C}}}(\mathcal{M})$. Assuming that noise is very small so that the distribution of $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ can be identified with the distribution over the tangent space $T_{\theta_{\mathcal{C}}}(\mathcal{C})$, we define the *covariance matrix* of the estimator $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ by

$$V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] = E\left[\left(\hat{\boldsymbol{\theta}}_{\mathcal{C}} - \boldsymbol{\theta}_{\mathcal{C}}\right)\left(\hat{\boldsymbol{\theta}}_{\mathcal{C}} - \boldsymbol{\theta}_{\mathcal{C}}\right)^{\top}\right]. \qquad (24)$$

This is an $n \times n$ singular matrix of rank $n - r$; its range is the tangent space $T_{\theta_{\mathcal{C}}}(\mathcal{C})$.

Since the covariance matrix is defined in the tangent space $T_{\theta_{\mathcal{C}}}(\mathcal{M})$, it suffices to introduce coordinates to $T_{\theta_{\mathcal{C}}}(\mathcal{M})$ *locally* rather than to the parameter space $\mathcal{M}$ globally.

Example 9: For the reduced orthographic model, let $\{\hat{\boldsymbol{s}}_{\alpha}\}$ and $\{\hat{\boldsymbol{R}}_{\kappa}\}$ be estimators of $\{\boldsymbol{s}_{\alpha}\}$ and $\{\boldsymbol{R}_{\kappa}\}$, respectively, under gauge $\mathcal{C}$. Since the shape vectors represent displacements, they constitute an additive group. Hence, the deviation from $\boldsymbol{s}_{\alpha}$ is measured by the *difference* $\Delta \boldsymbol{s}_{\alpha} = \hat{\boldsymbol{s}}_{\alpha} - \boldsymbol{s}_{\alpha}$. Rotations, on the other hand, constitute a multiplicative group. Hence, the deviation from $\boldsymbol{R}_{\kappa}$ is measured by the *quotient* $\hat{\boldsymbol{R}}_{\kappa} \boldsymbol{R}_{\kappa}^{\top}$, which is a small rotation. Let $\boldsymbol{l}_{\kappa}$ be its axis (unit vector) and $\Delta \Omega_{\kappa}$ the angle of rotation around it. If we put $\Delta \boldsymbol{\Omega}_{\kappa} = \Delta \Omega_{\kappa} \boldsymbol{l}_{\kappa}$, the estimator $\hat{\boldsymbol{R}}_{\kappa}$ is expressed to a first approximation in the form

$$\hat{\boldsymbol{R}}_{\kappa} = \boldsymbol{R}_{\kappa} + \Delta \boldsymbol{\Omega}_{\kappa} \times \boldsymbol{R}_{\kappa}. \qquad (25)$$

We can view $\{\Delta \boldsymbol{s}_{\alpha}\}$ and $\{\Delta \boldsymbol{\Omega}_{\kappa}\}$ as *local coordinates*[5] of the tangent space $T_{\theta_{\mathcal{C}}}(\mathcal{M})$. The covariance matrix of $\{\hat{\boldsymbol{s}}_{\alpha}\}$ and $\{\hat{\boldsymbol{R}}_{\kappa}\}$ is defined by

$$V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] = \begin{pmatrix} E\left[\Delta \boldsymbol{s}_1 \Delta \boldsymbol{s}_1^{\top}\right] & \cdots & E\left[\Delta \boldsymbol{s}_1 \Delta \boldsymbol{\Omega}_M^{\top}\right] \\ \vdots & & \vdots \\ E\left[\Delta \boldsymbol{\Omega}_M \Delta \boldsymbol{s}_1^{\top}\right] & \cdots & E\left[\Delta \boldsymbol{\Omega}_M \Delta \boldsymbol{\Omega}_M^{\top}\right] \end{pmatrix}.$$

$$(26)$$

X. GAUGE INVARIANTS

Recall that the indeterminacy of the solution is due to the overparameterization of the model and the loss of information during projection. Suppose there is a quantity that takes the same value everywhere on the leaf $\mathcal{M}_{\theta}$ associated with solution $\boldsymbol{\theta}$. This quantity is uniquely determined in spite of the indeterminacy of the solution.

Formally, a function $I: \mathcal{M} \to \mathcal{R}$ is *gauge invariant* if

$$I(g\boldsymbol{\theta}) = I(g'\boldsymbol{\theta}) \qquad (27)$$

---

[5]They correspond to the *Lie algebra* of the parameter space $\mathcal{M}$ if it is viewed as a Lie group. See footnote 1.

for all g and g' that belong to the same *connected component* of the gauge group $\mathcal{G}$. It follows that a function I is a gauge invariant *if and only if it is invariant to infinitesimal gauge transformations*. Let γ be an infinitesimal gauge transformation, and $\boldsymbol{D}(\cdot)$ its generator. Gauge invariance implies

$$\begin{aligned} I(\boldsymbol{\theta}) &= I(\gamma\boldsymbol{\theta}) = I(\boldsymbol{\theta} + \boldsymbol{D}(\boldsymbol{\theta}) + \cdots) \\ &= I(\boldsymbol{\theta}) + (\nabla_{\theta}I, \boldsymbol{D}(\boldsymbol{\theta})) + \cdots. \end{aligned} \qquad (28)$$

Let $\{\boldsymbol{D}_1(\cdot), \ldots, \boldsymbol{D}_r(\cdot)\}$ be an arbitrary basis of the linear space $\mathcal{D}_{\theta}(\mathcal{M})$ of gauge generators. The corresponding vectors $\boldsymbol{D}_1(\boldsymbol{\theta}), \ldots, \boldsymbol{D}_r(\boldsymbol{\theta})$ are the basis of the tangent space $T_{\theta}(\mathcal{M}_{\theta})$ to the leaf $\mathcal{M}_{\theta}$ associated with $\boldsymbol{\theta}$. Equation (28) implies that $(\nabla_{\theta}I, \boldsymbol{D}_i(\boldsymbol{\theta})) = 0, i = 1, \ldots, r$. Conversely, these differential equations generate a gauge invariant by integration. Thus, we have the following theorem (the superscript $\perp$ denotes orthogonal complement).

Theorem 1: A function $I: \mathcal{M} \to \mathcal{R}$ is a gauge invariant if and only if

$$\nabla_{\theta}I \in T_{\theta}(\mathcal{M}_{\theta})^{\perp} \qquad (29)$$

for all $\boldsymbol{\theta}$.

XI. EQUIVALENCE OF COVARIANCE MATRICES

Not only is the value of a gauge invariant unique, but its covariance is also uniquely obtainable given the covariance of the measured data. In contrast, the covariance of a quantity that depends on gauges cannot be uniquely specified: we have many possible "geometrically equivalent" covariances that correspond to the same measurement covariance. In this section, we derive a geometric equivalence relationship between covariance matrices. This permits us to test whether or not covariances of one parameter under different gauges correspond to the same underlying measurement covariance.

Let $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ and $\hat{\boldsymbol{\theta}}_{\mathcal{C}'}$ be estimators of $\boldsymbol{\theta}$ with respect to different gauges $\mathcal{C}$ and $\mathcal{C}'$, respectively, and let $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ and $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}]$ be their respective covariance matrices. We say that the covariance matrices $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ and $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}]$ are *geometrically equivalent* if the uncertainty of any gauge invariant is the same whichever covariance matrix we use. Let $I(\boldsymbol{\theta})$ be a gauge invariant. Its variance is to a first approximation

$$\begin{aligned} V[I] &= E\left[(\nabla_{\theta}I|_{\theta_{\mathcal{C}}}, \Delta\boldsymbol{\theta}_{\mathcal{C}})^2\right] \\ &= \left(\nabla_{\theta}I|_{\theta_{\mathcal{C}}}, E\left[\Delta\boldsymbol{\theta}_{\mathcal{C}} \Delta\boldsymbol{\theta}_{\mathcal{C}}^{\top}\right] \nabla_{\theta}I|_{\theta_{\mathcal{C}}}\right) \\ &= \left(\nabla_{\theta}I|_{\theta_{\mathcal{C}}}, V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] \nabla_{\theta}I|_{\theta_{\mathcal{C}}}\right) \end{aligned} \qquad (30)$$

where $\Delta\boldsymbol{\theta}_{\mathcal{C}} = \hat{\boldsymbol{\theta}}_{\mathcal{C}} - \boldsymbol{\theta}_{\mathcal{C}}$.

Let $\boldsymbol{\theta}_{\mathcal{C}}$ and $\boldsymbol{\theta}_{\mathcal{C}'}$ be the true values (the values we would have in the absence of noise) of $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ and $\hat{\boldsymbol{\theta}}_{\mathcal{C}'}$, respectively. Since $\boldsymbol{\theta}_{\mathcal{C}}$ and $\boldsymbol{\theta}_{\mathcal{C}'}$ both belong to the true leaf $\mathcal{M}_{\theta}$, there exists a gauge transformation $g \in \mathcal{G}$ such that $\boldsymbol{\theta}_{\mathcal{C}} = g\boldsymbol{\theta}_{\mathcal{C}'}$. This transformation of $\mathcal{M}$ induces a linear mapping from the tangent space $T_{\theta_{\mathcal{C}'}}(\mathcal{M})$ at $\boldsymbol{\theta}_{\mathcal{C}'}$ to the tangent space $T_{\theta_{\mathcal{C}}}(\mathcal{M})$ at $\boldsymbol{\theta}_{\mathcal{C}}$; each is described with respect to a local coordinate system. The mapping is symbolically denoted by $\partial\boldsymbol{\theta}_{\mathcal{C}}/\partial\boldsymbol{\theta}_{\mathcal{C}'}$ and called the *Jacobian matrix* from $T_{\theta_{\mathcal{C}'}}(\mathcal{M})$ to $T_{\theta_{\mathcal{C}}}(\mathcal{M})$. Then, $\hat{\boldsymbol{\theta}}_{\mathcal{C}'} = \boldsymbol{\theta}_{\mathcal{C}'} + \Delta\boldsymbol{\theta}_{\mathcal{C}'}$ is geometrically

equivalent to $\hat{\boldsymbol{\theta}}_{\mathcal{C}} = \boldsymbol{\theta}_{\mathcal{C}} + (\partial\boldsymbol{\theta}_{\mathcal{C}}/\partial\boldsymbol{\theta}_{\mathcal{C}'})\,\Delta\boldsymbol{\theta}_{\mathcal{C}'}$ to a first approximation. Hence,

$$V[I] = \left(\nabla_\theta I|_{\theta_c},\ \frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}\, V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}\right] \left(\frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}\right)^{\top} \nabla_\theta I|_{\theta_c} \right). \quad (31)$$

After substituting (31) from (30), Theorem 1 implies that $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ and $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}]$ are geometrically equivalent if and only if

$$\left(\boldsymbol{u},\ \left(V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] - \frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}\, V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}\right] \left(\frac{\partial\boldsymbol{\theta}_{\mathcal{C}}}{\partial\boldsymbol{\theta}_{\mathcal{C}'}}\right)^{\top} \right) \boldsymbol{u} \right) = 0 \quad (32)$$

for all $\boldsymbol{u} \in T_{\theta_c}(\mathcal{M}_\theta)^{\perp}$. This means that $T_{\theta_c}(\mathcal{M}_\theta)^{\perp}$ is contained in the null space of the symmetric matrix

$$V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}] - (\partial\boldsymbol{\theta}_{\mathcal{C}}/\partial\boldsymbol{\theta}_{\mathcal{C}'})V[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}](\partial\boldsymbol{\theta}_{\mathcal{C}}/\partial\boldsymbol{\theta}_{\mathcal{C}'})^{\top}.$$

In other words, its range is contained in $T_{\theta_c}(\mathcal{M}_\theta)$, in which all eigenvectors for nonzero eigenvalues lie.

Intuitively, this means that $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ and $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}]$ differ only in directions orthogonal to the leaf $\mathcal{M}_\theta$. Since the Jacobian matrix $\partial\boldsymbol{\theta}_{\mathcal{C}}/\partial\boldsymbol{\theta}_{\mathcal{C}'}$ is introduced merely for comparing them in the same local coordinate system, this is a symmetric relation between the gauges $\mathcal{C}$ and $\mathcal{C}'$. We introduce the notation

$$V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}] \equiv V[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}]\,\mathrm{mod}\,\mathcal{M}_\theta$$

for this relation and conclude as follows.

Theorem 2: Covariance matrices $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ and $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}]$ for different gauges are geometrically equivalent if and only if

$$V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}] \equiv V[\hat{\boldsymbol{\theta}}_{\mathcal{C}'}]\,\mathrm{mod}\ \mathcal{M}_\theta. \quad (33)$$

Example 10: For the reduced orthographic model, the parameter space $\mathcal{M}$ is locally parameterized by $\{\{\Delta\boldsymbol{s}_\alpha\}, \{\Delta\boldsymbol{\Omega}_\kappa\}\}$. The Jacobian matrix $\partial\boldsymbol{\theta}'/\partial\boldsymbol{\theta}$ is defined as a linear mapping from $\{\{\Delta\boldsymbol{s}_\alpha\}, \{\Delta\boldsymbol{\Omega}_\kappa\}\}$ at $\{\{\boldsymbol{s}_\alpha\}, \{\boldsymbol{R}_\kappa\}\}$ to $\{\{\Delta\boldsymbol{s}'_\alpha\}, \{\Delta\boldsymbol{\Omega}'_\kappa\}\}$ at $\{\boldsymbol{s}'_\alpha\}, \{\boldsymbol{R}'_\kappa\}\}$ given the gauge transformation $\{\boldsymbol{R}, \boldsymbol{t}\}$ that maps $\{\{\boldsymbol{s}_\alpha\}, \{\boldsymbol{R}_\kappa\}\}$ to $\{\{\boldsymbol{s}'_\alpha\}, \{\boldsymbol{R}'_\kappa\}\}$. From (15), we obtain

$$\Delta\boldsymbol{s}'_\alpha = \boldsymbol{R}^{\top}\Delta\boldsymbol{s}_\alpha \qquad \Delta\boldsymbol{\Omega}'_\kappa = \Delta\boldsymbol{\Omega}_\kappa. \quad (34)$$

The latter results from

$$\begin{aligned}
\boldsymbol{R}'_\kappa + \Delta\boldsymbol{\Omega}'_\kappa \times \boldsymbol{R}'_\kappa &= (\boldsymbol{R}_\kappa + \Delta\boldsymbol{\Omega}_\kappa \times \boldsymbol{R}_\kappa)\boldsymbol{R} \\
&= \boldsymbol{R}_\kappa\boldsymbol{R} + \Delta\boldsymbol{\Omega}_\kappa \times \boldsymbol{R}_\kappa\boldsymbol{R}' \\
&= \boldsymbol{R}'_\kappa + \Delta\boldsymbol{\Omega}_\kappa \times \boldsymbol{R}'_\kappa.
\end{aligned}$$

The Jacobian matrix has the form

$$\frac{\partial\boldsymbol{\theta}'}{\partial\boldsymbol{\theta}} = \mathrm{diag}(\boldsymbol{R}^{\top}, \ldots, \boldsymbol{R}^{\top}, \boldsymbol{I}, \ldots, \boldsymbol{I}) \quad (35)$$

where $\mathrm{diag}(\boldsymbol{A}, \ldots, \boldsymbol{B})$ denotes the block-diagonal matrix with $\boldsymbol{A}, \ldots, \boldsymbol{B}$ as its diagonal blocks in that order.

XII. EQUIVALENCE OF ESTIMATORS

Assume that we have obtained an estimate of the solution parameter under a particular gauge. We would like to identify which perturbations of it are geometrically equivalent, i.e., which perturbations would correspond to the same perturbation of the image data. In this section, we define an oblique projection onto the gauge manifold and show that all perturbations that project onto identical points are geometrically equivalent.

Let $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ be an estimator of $\boldsymbol{\theta}$ with respect to a gauge $\mathcal{C}$. To a first approximation, the deviation $\Delta\boldsymbol{\theta}_{\mathcal{C}} = \hat{\boldsymbol{\theta}}_{\mathcal{C}} - \boldsymbol{\theta}_{\mathcal{C}}$ can be identified with an element in $T_{\theta_c}(\mathcal{C})$. Let $\Delta\boldsymbol{\theta} \in T_{\theta_c}(\mathcal{M})$ be an

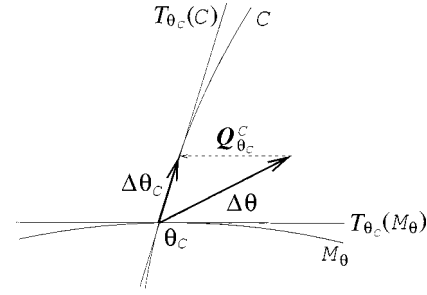Fig. 1.   Oblique projection onto the tangent space to $\mathcal{C}$.

arbitrary vector. To a first approximation, $\boldsymbol{\theta}_{\mathcal{C}} + \Delta\boldsymbol{\theta}$ is geometrically equivalent to $\hat{\boldsymbol{\theta}}_{\mathcal{C}} = \boldsymbol{\theta}_{\mathcal{C}} + \Delta\boldsymbol{\theta}_{\mathcal{C}}$ if and only if $\Delta\boldsymbol{\theta} - \Delta\boldsymbol{\theta}_{\mathcal{C}} \in T_{\theta_c}(\mathcal{M}_\theta)$ (Fig. 1). Since $\{\boldsymbol{D}_1(\boldsymbol{\theta}_{\mathcal{C}}), \ldots, \boldsymbol{D}_r(\boldsymbol{\theta}_{\mathcal{C}})\}$ are the basis of $T_{\theta_c}(\mathcal{M}_\theta)$, this condition is equivalent to the existence of r numbers $x_1, \ldots, x_r$ such that $\Delta\boldsymbol{\theta}_{\mathcal{C}} = \Delta\boldsymbol{\theta} + \sum_{i=1}^{r} x_i\boldsymbol{D}_i(\boldsymbol{\theta}_{\mathcal{C}})$ or

$$\Delta\boldsymbol{\theta}_{\mathcal{C}} = \Delta\boldsymbol{\theta} + \boldsymbol{U}_{\theta_c}\boldsymbol{x} \quad (36)$$

where we let $\boldsymbol{x} = (x_1, \ldots, x_r)^{\top}$ and define

$$\boldsymbol{U}_{\theta_c} = (\ \boldsymbol{D}_1(\boldsymbol{\theta}_{\mathcal{C}}) \quad \cdots \quad \boldsymbol{D}_r(\boldsymbol{\theta}_{\mathcal{C}})\ ). \quad (37)$$

If the gauge $\mathcal{C}$ is defined by r equations $c_1(\boldsymbol{\theta}) = 0, \ldots, c_r(\boldsymbol{\theta}) = 0$, the tangent space $T_{\theta_c}(\mathcal{C})$ is the orthogonal complement of the linear space spanned by $\{\nabla_\theta c_1|_{\theta_c}, \ldots \nabla_\theta c_r|_{\theta_c}\}$. It follows that $(\nabla_\theta c_i|_{\theta_c}, \Delta\boldsymbol{\theta}_{\mathcal{C}}) = 0$ for $i = 1, \ldots, r$. From (36), we have

$$\boldsymbol{V}_{\theta_c}^{\top}\Delta\boldsymbol{\theta} + \boldsymbol{V}_{\theta_c}^{\top}\boldsymbol{U}_{\theta_c}\boldsymbol{x} = \boldsymbol{0} \quad (38)$$

where we have defined

$$\boldsymbol{V}_{\theta_c} = (\ \nabla_\theta c_1|_{\theta_c} \quad \cdots \quad \nabla_\theta c_r|_{\theta_c}\ ). \quad (39)$$

Eliminating $\boldsymbol{x}$ from (36) and (38), we obtain

$$\Delta\boldsymbol{\theta}_{\mathcal{C}} = \Delta\boldsymbol{\theta} - \boldsymbol{U}_{\theta_c}\left(\boldsymbol{V}_{\theta_c}^{\top}\boldsymbol{U}_{\theta_c}\right)^{-1}\boldsymbol{V}_{\theta_c}^{\top}\Delta\boldsymbol{\theta} = \boldsymbol{Q}_{\theta_c}^{\mathcal{C}}\Delta\boldsymbol{\theta} \quad (40)$$

where

$$\boldsymbol{Q}_{\theta_c}^{\mathcal{C}} = \boldsymbol{I} - \boldsymbol{U}_{\theta_c}(\boldsymbol{V}_{\theta_c}^{\top}\boldsymbol{U}_{\theta_c})^{-1}\boldsymbol{V}_{\theta_c}^{\top} \quad (41)$$

which defines an (oblique) projection onto $T_{\theta_c}(\mathcal{C})$ along $T_{\theta_c}(\mathcal{M}_\theta)$ (Fig. 1). Thus, we have the following theorem.

Theorem 3: An estimator $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ is geometrically equivalent to $\boldsymbol{\theta}_{\mathcal{C}} + \Delta\boldsymbol{\theta}$ to a first approximation if and only if

$$\hat{\boldsymbol{\theta}}_{\mathcal{C}} = \boldsymbol{\theta}_{\mathcal{C}} + \boldsymbol{Q}_{\theta_c}^{\mathcal{C}}\Delta\boldsymbol{\theta}. \quad (42)$$

Example 11: For the reduced orthographic model, the standard gauge (21) is expressed in the local coordinates $\{\Delta\boldsymbol{s}_\alpha\}$ and $\{\Delta\boldsymbol{\Omega}_\kappa\}$ in the form

$$\sum_{\alpha=1}^{N}\Delta\boldsymbol{s}_\alpha = \boldsymbol{0} \qquad \Delta\boldsymbol{\Omega}_1 = \boldsymbol{0}. \quad (43)$$

This means that the six basis vectors of the tangent space to the standard gauge manifold $\mathcal{C}$ are the columns of the $3(M+N)\times 6$ matrix

$$\boldsymbol{V}_\theta = \begin{pmatrix} \boldsymbol{I} & \cdots & \boldsymbol{I} & \boldsymbol{O} & \boldsymbol{O} & \cdots\boldsymbol{O} \\ \boldsymbol{O} & \cdots & \boldsymbol{O} & \boldsymbol{I} & \boldsymbol{O} & \cdots\boldsymbol{O} \end{pmatrix}^{\top}. \quad (44)$$

Combining (17) and (25), we can write $\hat{R}_\kappa = R_\kappa + D(R_\kappa)$ as $\Delta\Omega_\kappa \times R_\kappa = -R_\kappa \times \Delta\Omega$, which is equivalent to $\Delta\Omega_\kappa = R_\kappa \Delta\Omega$. Hence, the gauge generator is expressed in terms of the local coordinates $\{\Delta s_\alpha\}$ and $\{\Delta\Omega_\kappa\}$ in the form

$$D(\boldsymbol{\theta}) = U_\theta \begin{pmatrix} \Delta\Omega \\ \Delta t \end{pmatrix}$$

$$U_\theta = \begin{pmatrix} -s_1 \times I & \cdots & -s_N \times I & R_1^\top & \cdots & R_M^\top \\ -I & \cdots & -I & O & \cdots & O \end{pmatrix}^\top. \quad (45)$$

If $\{s_\alpha\}$ and $\{R_\kappa\}$ satisfy the standard gauge (21), we have

$$Q_\theta^C = I - \begin{pmatrix} I/N & \cdots & I/N & s_1 \times I & O & \cdots & O \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ I/N & \cdots & I/N & s_N \times I & O & \cdots & O \\ O & \cdots & O & I & O & \cdots & O \\ O & \cdots & O & R_2 & O & \cdots & O \\ \vdots & & \vdots & \vdots & \vdots & \cdots & \vdots \\ O & \cdots & O & R_M & O & \cdots & O \end{pmatrix}. \quad (46)$$

It follows that the standard gauge (21) is enforced to the local coordinates $\{\Delta s_\alpha\}$ and $\{\Delta\Omega_\kappa\}$ in the form

$$\Delta s_\alpha|_S \leftarrow \Delta s_\alpha - \frac{1}{N}\sum_{\beta=1}^N \Delta s_\beta - s_\alpha \times \Delta\Omega_1$$

$$\Delta\Omega_\kappa|_S \leftarrow \begin{cases} 0, & \kappa = 1 \\ \Delta\Omega_\kappa - R_\kappa \Delta\Omega_1, & \kappa = 2, \ldots, M \end{cases} \quad (47)$$

where $|_S$ denotes quantities after the standard gauge is enforced. The intuitive meaning is that (43) are enforced by subtracting from Δs_α the centroid of $\{\Delta s_\beta\}$ and rotating the frame back by $-\Delta\Omega_1$.

XIII. COVARIANCE TRANSFORMATIONS

Our oblique projection provides us with a convenient way to transform covariances between different gauges. Let $\hat{\boldsymbol{\theta}}_C$ be an estimator of $\boldsymbol{\theta}$ with respect to gauge C, and $V[\hat{\boldsymbol{\theta}}_C]$ its covariance matrix. Let $\hat{\boldsymbol{\theta}}_{C'}$ be a geometrically equivalent estimator of $\boldsymbol{\theta}$ with respect to another gauge C', and write

$$\hat{\boldsymbol{\theta}}_C = \boldsymbol{\theta}_C + \Delta\boldsymbol{\theta}_C \qquad \hat{\boldsymbol{\theta}}_{C'} = \boldsymbol{\theta}_{C'} + \Delta\boldsymbol{\theta}_{C'}. \quad (48)$$

The former is geometrically equivalent to $\boldsymbol{\theta}_{C'} + (\partial\boldsymbol{\theta}_{C'}/\partial\boldsymbol{\theta}_C)\Delta\boldsymbol{\theta}_C$. According to Theorem 3, the estimators $\hat{\boldsymbol{\theta}}_C$ and $\hat{\boldsymbol{\theta}}_{C'}$ are geometrically equivalent if and only if

$$\Delta\boldsymbol{\theta}_{C'} = Q_{\theta_{C'}}^{C'}(\partial\boldsymbol{\theta}_{C'}/\partial\boldsymbol{\theta}_C)\Delta\boldsymbol{\theta}_C.$$

From this, we obtain the following rule of transformation of the covariance matrix induced by a change of the gauge.

Theorem 4: Let $V[\hat{\boldsymbol{\theta}}_C]$ be the covariance matrix of estimator $\hat{\boldsymbol{\theta}}_C$ under gauge C. For another gauge C', the covariance matrix of the corresponding estimator $\hat{\boldsymbol{\theta}}_{C'}$ is given by

$$V[\hat{\boldsymbol{\theta}}_{C'}] = Q_{\theta_{C'}}^{C'} \frac{\partial\boldsymbol{\theta}_{C'}}{\partial\boldsymbol{\theta}_C} V[\hat{\boldsymbol{\theta}}_C]\left(\frac{\partial\boldsymbol{\theta}_{C'}}{\partial\boldsymbol{\theta}_C}\right)^\top Q_{\theta_{C'}}^{C'\top}. \quad (49)$$

XIV. UNBIASED ESTIMATORS

We have so far assumed that an estimator $\hat{\boldsymbol{\theta}}_C$ under gauge C takes the "true value" $\boldsymbol{\theta}_C \in \mathcal{M}_\theta \cap C$ in the absence of noise. We now strengthen this condition: we require the estimator $\hat{\boldsymbol{\theta}}_C$ to be "unbiased." The formal definition is as follows.

Let $p(\{\boldsymbol{p}_\alpha\}; \boldsymbol{\theta})$ be the probability density of the data $\{\boldsymbol{p}_\alpha\}$ parameterized by $\boldsymbol{\theta}$. We define its *score* l_θ by

$$l_\theta = \nabla_\theta \log p. \quad (50)$$

The *Fisher information matrix* is defined by

$$J_\theta = E[l_\theta l_\theta^\top] \quad (51)$$

which is a symmetric positive semidefinite matrix. Note that $p(\{\boldsymbol{p}_\alpha\}; \boldsymbol{\theta})$ is a gauge invariant and hence $\log p$ is also a gauge invariant. According to Theorem 1, we see that

$$l_\theta = \nabla_\theta \log p \in T_\theta(\mathcal{M}_\theta)^\perp.$$

It follows that the Fisher information matrix J_{θ_C} has rank $n - r$, having the range $T_\theta(\mathcal{M}_\theta)^\perp$ and the null space $T_\theta(\mathcal{M}_\theta)$.

An estimator $\hat{\boldsymbol{\theta}}$ under gauge C is *unbiased* if

$$E[\hat{\boldsymbol{\theta}}_C - \boldsymbol{\theta}_C] = 0 \quad (52)$$

where we are assuming that the deviation $\hat{\boldsymbol{\theta}}_C - \boldsymbol{\theta}_C$ is so small that it can be identified with an element in the tangent space $T_{\theta_C}(C)$. Equation (52) is rewritten as

$$\int \cdots \int (\hat{\boldsymbol{\theta}}_C - \boldsymbol{\theta}_C) p(\{\boldsymbol{p}_\alpha\}; \boldsymbol{\theta}_C)\, d\boldsymbol{p}_1 \cdots \boldsymbol{p}_N = 0. \quad (53)$$

This should hold for any $\boldsymbol{\theta}_C \in C$ and hence for an arbitrary infinitesimal perturbation $\boldsymbol{\theta}_C \to \boldsymbol{\theta}_C + \delta\boldsymbol{\theta}_C$ for $\delta\boldsymbol{\theta}_C \in T_{\theta_C}(C)$. After some manipulations using the identity $\nabla_\theta p = p\nabla_\theta \log p$, we conclude that

$$\left(E\left[(\hat{\boldsymbol{\theta}}_C - \boldsymbol{\theta}_C)l_{\theta_C}^\top\right] - I\right)\delta\boldsymbol{\theta}_C = 0 \quad (54)$$

should hold for an arbitrary vector $\delta\boldsymbol{\theta}_C \in T_{\theta_C}(C)$ and hence for $\delta\boldsymbol{\theta}_C = Q_{\theta_C}^C \boldsymbol{x} \in T_{\theta_C}(C)$ for an arbitrary vector $\boldsymbol{x}$. It follows that

$$\left(E\left[(\hat{\boldsymbol{\theta}}_C - \boldsymbol{\theta}_C)l_{\theta_C}^\top\right] - I\right)Q_{\theta_C}^C = O \quad (55)$$

which can be rewritten as

$$E\left[\left(\hat{\boldsymbol{\theta}}_C - \boldsymbol{\theta}_C\right)l_{\theta_C}^\top\right] = Q_{\theta_C}^C \quad (56)$$

if we note that

$$l_{\theta_C}^\top Q_{\theta_C}^C = (Q_{\theta_C}^{C\top} l_{\theta_C})^\top = l_{\theta_C}^\top.$$

In fact, we see that for an arbitrary vector $\boldsymbol{x}$

$$\begin{aligned}\left(l_{\theta_C} - Q_{\theta_C}^{C\top} l_{\theta_C}, \boldsymbol{x}\right) &= (l_{\theta_C}, \boldsymbol{x}) - \left(Q_{\theta_C}^{C\top} l_{\theta_C}, \boldsymbol{x}\right) \\ &= (l_{\theta_C}, \boldsymbol{x}) - \left(l_{\theta_C}, Q_{\theta_C}^C \boldsymbol{x}\right) \\ &= \left(l_{\theta_C}, \boldsymbol{x} - Q_{\theta_C}^C \boldsymbol{x}\right) = 0 \quad (57)\end{aligned}$$

because $l_\theta \in T_\theta(\mathcal{M}_\theta)^\perp$ and $\boldsymbol{x} - Q_{\theta_C}^C \boldsymbol{x} \in T_\theta(\mathcal{M}_\theta)$.

XV. CRAMER–RAO INEQUALITY

We would like a bound on our parameter estimation accuracy that takes into account the gauge freedom of the parameters. To

achieve this, we derive a generalized form of the *Cramer–Rao lower bound* [6], [7]. Using identity (56), we obtain

$$E\left[\begin{pmatrix} \hat{\boldsymbol{\theta}}_{\mathcal{C}} - \boldsymbol{\theta}_{\mathcal{C}} \\ l_{\theta_{\mathcal{C}}} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\theta}}_{\mathcal{C}} - \boldsymbol{\theta}_{\mathcal{C}} \\ l_{\theta_{\mathcal{C}}} \end{pmatrix}^{\top}\right]$$

$$= \begin{pmatrix} V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] & \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \\ \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top} & \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top} \boldsymbol{J}_{\theta_{\mathcal{C}}} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \end{pmatrix} \quad (58)$$

which is positive semidefinite by construction. Hence, the following matrix should also be positive semidefinite:

$$\begin{pmatrix} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} & -\boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} \\ \boldsymbol{O} & \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} \end{pmatrix} \begin{pmatrix} V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] & \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \\ \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top} & \boldsymbol{J}_{\theta_{\mathcal{C}}} \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top} & \boldsymbol{O} \\ -\boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top} & \boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top} \end{pmatrix}. \quad (59)$$

Here, the superscript "$-$" denotes (Moore–Penrose) generalized inverse. Note that $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ has the range $T_{\theta_{\mathcal{C}}}(\mathcal{C})$ and hence $\boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}] \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top} = V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$. Using the identity $\boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} \boldsymbol{J}_{\theta_{\mathcal{C}}} \boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} = \boldsymbol{J}_{\theta_{\mathcal{C}}}^{-}$, we see that the above matrix equals

$$\begin{pmatrix} V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] - \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top} \end{pmatrix}. \quad (60)$$

Since this is positive semidefinite, so should $\boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top}$ be. Hence we have the following.

Theorem 5: For an estimator $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ under gauge $\mathcal{C}$, the following inequality holds:

$$V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] \succ \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \boldsymbol{J}_{\theta_{\mathcal{C}}}^{-} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top}. \quad (61)$$

Here, the inequality $\boldsymbol{A} \succ \boldsymbol{B}$ for symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite. We call the right-hand side of (61) the *Cramer–Rao lower bound under gauge* $\mathcal{C}$.

XVI. NORMAL FORM

So far, the covariance matrix of an estimator of $\boldsymbol{\theta}$ has been defined for a particular gauge. We now show that it can be defined *independently* of gauges. Let $\boldsymbol{\theta} \in \mathcal{M}_{\theta}$ be an arbitrary true value, and $\hat{\boldsymbol{\theta}}$ its estimator under an arbitrary gauge. Since $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$ may have very different values because of the gauge freedom, we cannot compare them directly. So, we apply an appropriate gauge transformation g such that $g\hat{\boldsymbol{\theta}}$ becomes the "closest" to $\boldsymbol{\theta}$ and then evaluate $E[(g\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(g\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^{\top}]$. The formal description is as follows.

Define a gauge manifold $\mathcal{C}^{*}$ such that i) it passes through $\boldsymbol{\theta}$ and ii) the tangent space $T_{\theta}(\mathcal{C}^{*})$ to $\mathcal{C}^{*}$ coincides with the *orthogonal complement* $T_{\theta}(\mathcal{M}_{\theta})^{\perp}$ of $T_{\theta}(\mathcal{M}_{\theta})$. We call the covariance matrix $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}^{*}}]$ of the estimator $\hat{\boldsymbol{\theta}}_{\mathcal{C}^{*}}$ under that gauge the *normal form* and denote it by $V^{*}[\hat{\boldsymbol{\theta}}]$ (Fig. 2). The normal form is defined for *all* values that belong to the leaf $\mathcal{M}_{\theta}$ of true values.

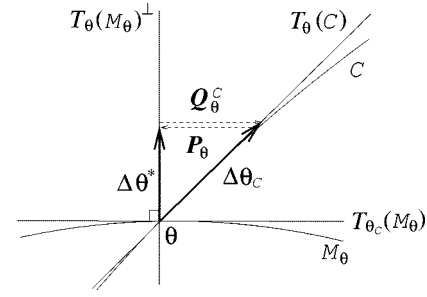Fig. 2.   Orthogonal projection from the tangent space to $\mathcal{C}$.

Let $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ be an estimator equivalent to $\hat{\boldsymbol{\theta}}_{\mathcal{C}^{*}}$ defined under another gauge $\mathcal{C}$ that also passes through $\boldsymbol{\theta}$. Let $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ be its covariance matrix. According to Theorem 4, the normal form $V^{*}[\hat{\boldsymbol{\theta}}]$ is given by

$$V^{*}[\hat{\boldsymbol{\theta}}] = \boldsymbol{P}_{\theta} V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] \boldsymbol{P}_{\theta}^{\top} \quad (62)$$

where $\boldsymbol{P}_{\theta}$ denotes the (orthogonal) projection onto $T_{\theta}(\mathcal{M}_{\theta})^{\perp}$ (Fig. 2). It has the expression

$$\boldsymbol{P}_{\theta} = \boldsymbol{I} - \boldsymbol{U}_{\theta} \left(\boldsymbol{U}_{\theta}^{\top} \boldsymbol{U}_{\theta}\right)^{-1} \boldsymbol{U}_{\theta}^{\top} \quad (63)$$

where $\boldsymbol{U}_{\theta}$ is the $n \times n$ matrix consisting of $\boldsymbol{D}_{1}(\boldsymbol{\theta}), \ldots, \boldsymbol{D}_{r}(\boldsymbol{\theta})$ as its columns (see (37)). Conversely, $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ is given in terms of its normal form $V^{*}[\hat{\boldsymbol{\theta}}]$ in the form

$$V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] = \boldsymbol{Q}_{\theta}^{\mathcal{C}} V^{*}[\hat{\boldsymbol{\theta}}] \boldsymbol{Q}_{\theta}^{\mathcal{C}\top}. \quad (64)$$

In general, we have the following.

Theorem 6: Let $\hat{\boldsymbol{\theta}}_{\mathcal{C}}$ be an estimator defined under a gauge $\mathcal{C}$, and $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ its covariance matrix. Its normal form at an arbitrary point $\boldsymbol{\theta} \in \mathcal{M}_{\theta}$ is

$$V^{*}[\hat{\boldsymbol{\theta}}] = \boldsymbol{P}_{\theta} \frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\theta}_{\mathcal{C}}} V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] \left(\frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\theta}_{\mathcal{C}}}\right)^{\top} \boldsymbol{P}_{\theta}^{\top}. \quad (65)$$

Conversely, $V[\hat{\boldsymbol{\theta}}_{\mathcal{C}}]$ is given in terms of its normal form at $\boldsymbol{\theta} \in \mathcal{M}_{\theta}$ in the form

$$V\left[\hat{\boldsymbol{\theta}}_{\mathcal{C}}\right] = \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}} \frac{\partial \boldsymbol{\theta}_{\mathcal{C}}}{\partial \boldsymbol{\theta}} V^{*}[\hat{\boldsymbol{\theta}}] \left(\frac{\partial \boldsymbol{\theta}_{\mathcal{C}}}{\partial \boldsymbol{\theta}}\right)^{\top} \boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}\top}. \quad (66)$$

In sum, the normal form and an arbitrary form can be transformed from one to the other via the orthogonal projection $\boldsymbol{P}_{\theta}$ (see (63)) and the oblique projection $\boldsymbol{Q}_{\theta_{\mathcal{C}}}^{\mathcal{C}}$ (see (41)). They are computed from the matrix $\boldsymbol{U}_{\theta_{\mathcal{C}}}$ (see (37)) and the matrix $\boldsymbol{V}_{\theta_{\mathcal{C}}}$ (see (39)). The former is defined by the gauge transformation $\mathcal{G}$ alone via (16); the latter by the gauge $\mathcal{C}$ given by (20) alone.

In practice, however, we can automatically obtain the normal form if we adopt the "free-gauge approach," which we will describe soon. This is based on the following observation. Recall that the Cramer–Rao lower bound given by (61) is defined for every true value $\boldsymbol{\theta}$ that belongs to the true leaf $\mathcal{M}_{\theta}$ for any given gauge $\mathcal{C}$. If the gauge $\mathcal{C}$ happens to be such that $T_{\theta}(\mathcal{C}) = T_{\theta}(\mathcal{M}_{\theta})^{\perp}$, (61) can be read as giving a lower bound on the covariance matrix of $\hat{\boldsymbol{\theta}}$ in the normal form. But if $T_{\theta}(\mathcal{C}) =$

$T_\theta(\mathcal{M}_\theta)^\perp$, then $\boldsymbol{Q}_{\theta_C}^{\mathcal{C}} = \boldsymbol{P}_\theta$ and the Fisher information matrix $\boldsymbol{J}_\theta$ has the range $T_\theta(\mathcal{M}_\theta)^\perp$. Thus, we obtain the following Cramer–Rao inequality.

Theorem 7: For any estimator $\hat{\boldsymbol{\theta}}$, the following inequality holds for the normal form of its covariance matrix:

$$V^*[\hat{\boldsymbol{\theta}}] \succ \boldsymbol{J}_\theta^-. \qquad (67)$$

XVII. NUMERICAL OPTIMIZATION

The question remains as to how to numerically obtain the optimal estimator under a given gauge. Newton-type optimization is attractive for its quadratic convergence property, but at the same time the solution should satisfy the required gauge condition. In this and the next section, we consider various Newton-type optimization methods that impose gauge conditions.

The function J given by (13) has the following Taylor expansion in the neighborhood of the true value $\overline{\boldsymbol{\theta}}_C$ under gauge $\mathcal{C}$:

$$J\left(\overline{\boldsymbol{\theta}}_{\mathcal{C}} + \Delta\boldsymbol{\theta}\right) = \overline{J} + (\nabla_\theta \overline{J}, \Delta\boldsymbol{\theta}) + \tfrac{1}{2}\left(\Delta\boldsymbol{\theta}, \nabla_\theta^2 \overline{J} \Delta\boldsymbol{\theta}\right) + \cdots. \qquad (68)$$

The bars in the expressions $\overline{J}$, $\nabla_\theta \overline{J}$, and $\nabla_\theta^2 \overline{J}$ mean that they are evaluated at the true value $\boldsymbol{\theta} = \overline{\boldsymbol{\theta}}_C$. Here, $\Delta\boldsymbol{\theta}$ is identified with an element of the tangent space $T_{\overline{\theta}_C}(\mathcal{M})$ at $\overline{\boldsymbol{\theta}}_C$ and hence described in terms of the *local coordinates* of $T_{\overline{\theta}_C}(\mathcal{M})$ (e.g., the vector $\Delta\boldsymbol{\Omega}$ for rotation $\boldsymbol{R}$; see Example 9).

Ignoring terms of order $O(\epsilon^3)$ and higher, differentiating (68) with respect to $\Delta\boldsymbol{\theta}$, letting the result be zero, and substituting an estimate $\boldsymbol{\theta}_C$ for the true value $\overline{\boldsymbol{\theta}}_C$, we obtain

$$\nabla_\theta J + \nabla_\theta^2 J \Delta\boldsymbol{\theta} = \boldsymbol{0}. \qquad (69)$$

Since the rank of the Hessian $\nabla_\theta^2 J$ is $n - r$, this equation has infinitely many solutions. Here, two alternatives exist. One is to combine (69) with the r equations

$$(\nabla_\theta c_1|_\theta, \Delta\boldsymbol{\theta}) = 0, \ldots, (\nabla_\theta c_r|_\theta, \Delta\boldsymbol{\theta}) = 0 \qquad (70)$$

which constrain $\Delta\boldsymbol{\theta}$ to be in the tangent space $T_\theta(\mathcal{C})$. The resulting simultaneous linear equations can determine $\Delta\boldsymbol{\theta}$ uniquely. An alternative method is to compute an arbitrary solution of (69) first. For example, we can choose the one that has the minimum norm $\|\Delta\boldsymbol{\theta}\|$. The solution is obtained by using the (Moore–Penrose) generalized inverse. Then, we impose the gauge $\mathcal{C}$ by replacing $\Delta\boldsymbol{\theta}$ by $\boldsymbol{Q}_\theta^{\mathcal{C}} \Delta\boldsymbol{\theta}$ according to Theorem 3 (see (47)). Thus, we have

$$\Delta\boldsymbol{\theta} = -\boldsymbol{Q}_\theta^{\mathcal{C}} \nabla_\theta^2 J^- \nabla_\theta J. \qquad (71)$$

The subscript "$n - r$" means that the smallest r eigenvalues are replaced by zeros [6].

Whichever method we use, the parameter $\boldsymbol{\theta}$ is updated in the form

$$\boldsymbol{\theta} \leftarrow \exp(\Delta\boldsymbol{\theta})\boldsymbol{\theta} \qquad (72)$$
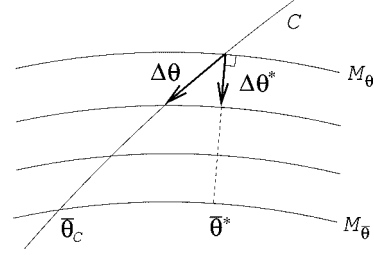

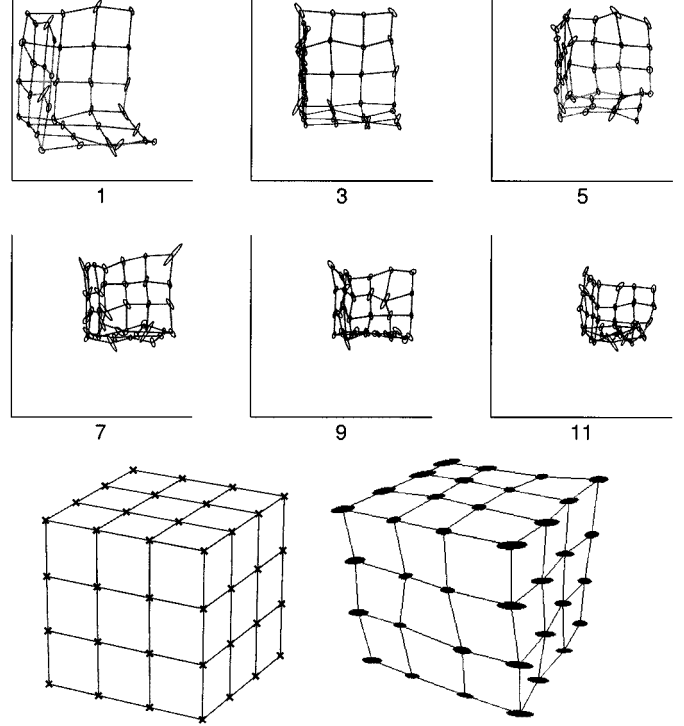Fig. 3. The gauge approach versus the free-gauge approach.

Fig. 4. A decimated sequence of synthetic images with uncertainty ellipses. On bottom left is the true object shape; on bottom right is an optimal 3-D reconstruction with uncertainty ellipsoids.

where $\exp(\cdot)$ is the *exponential mapping* from the tangent space $T_{\overline{\theta}_C}(\mathcal{M})$ to $\mathcal{M}$ to the parameter space $\mathcal{M}$ itself[6]: it is defined in such a way that $\exp(\Delta\boldsymbol{\theta})\boldsymbol{\theta}$ is in $\mathcal{M}$ and equal to $\boldsymbol{\theta} + \Delta\boldsymbol{\theta}$ to a first approximation. Application of this mapping is necessary because $\boldsymbol{\theta} + \Delta\boldsymbol{\theta}$ may not be in $\mathcal{M}$ exactly. In addition, we further need a higher order adjustment to enforce the gauge $\mathcal{C}$ exactly, because (69) is a first-order approximation and hence $\boldsymbol{\theta}$ may not be in $\mathcal{C}$ exactly.

Example 12: For translation, the exponential mapping acts in the form

$$\exp(\Delta\boldsymbol{t})\boldsymbol{t} = \boldsymbol{t} + \Delta\boldsymbol{t}. \qquad (73)$$

For rotation, we have

$$\exp(\Delta\boldsymbol{\Omega})\boldsymbol{R}$$
$$= \left(\cos\Delta\Omega \boldsymbol{I} + (1 - \cos\Delta\Omega)\boldsymbol{l}\boldsymbol{l}^\top + \sin\Delta\Omega \boldsymbol{l} \times \boldsymbol{I}\right)\boldsymbol{R} \qquad (74)$$

---

[6]Mathematically, this is a mapping from the Lie algebra of a Lie group to the Lie group itself.
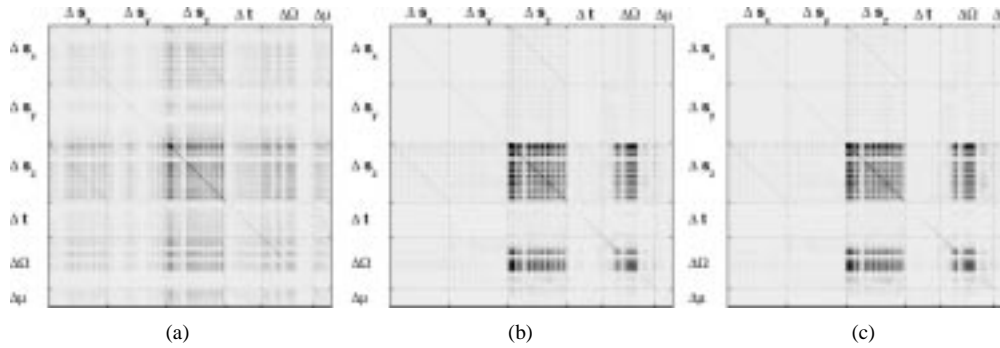
Fig. 5. (a) Predicted normal covariance. (b) Covariance description for the standard gauge. (c) Monte Carlo simulation for the standard gauge.

where $\Delta\Omega = \|\Delta\Omega\|$ and $l = \Delta\Omega/\|\Delta\Omega\|$. The right-hand side is equal to $R + \Delta\Omega \times R$ to a first approximation (see (25)). The matrix $\exp(\Delta\Omega)$ is a rotation by angle $\|\Delta\Omega\|$ around axis $\Delta\Omega$. The above expression is known as the *Rodrigues formula* [4].

XVIII. FREE-GAUGE APPROACH

If we carefully examine the origin of the indeterminacy of the solution of (69), we find that gauges need not necessarily be imposed at first. Instead, we may choose an arbitrary solution $\Delta\theta$ and update θ by (72) until the solution converges. The gauge $\mathcal{C}$ is imposed at this stage. We call this strategy, first discussed by Triggs [21], the *free-gauge approach*. Evidently, computationally the most convenient of the solutions of (69) is the minimum-norm solution (Fig. 3)

$$\Delta\theta^* = -\nabla_\theta^2 J^- \nabla_\theta J. \tag{75}$$

It is easily seen that the maximum-likelihood estimator attains the Cramer–Rao lower bound to a first approximation under Gaussian noise [6]. Hence, we can approximately evaluate the Cramer–Rao lower bound by evaluating J_θ at the computed estimate $\hat{\theta}$. If we ignore terms of order $O(\epsilon^3)$ and higher, the Fisher information matrix takes the form

$$J_\theta|_{\hat{\theta}} = \frac{1}{2}\nabla_\theta^2 J|_{\hat{\theta}}. \tag{76}$$

It follows that the free-gauge approach with the choice of (75) allows us to obtain not only a maximum-likelihood solution $\hat{\theta}$ but also its covariance matrix in the normal form as follows:

$$V^*[\hat{\theta}] = -2\left(\nabla_\theta^2 J|_{\hat{\theta}}\right)^-. \tag{77}$$

The free-gauge approach may be advantageous when the gauge manifold $\mathcal{C}$ is "nearly parallel" to the leaves $\mathcal{M}_\theta$, $\theta \in \mathcal{M}$, since the numerical solution must travel along a long path to reach the solution leaf $\mathcal{M}_{\bar{\theta}}$ if the trajectory is constrained to be in $\mathcal{C}$. The free-gauge approach allows trajectories to intersect the individual leaves $\mathcal{M}_\theta$, $\theta \in \mathcal{M}$, "orthogonally," so the true leaf $\mathcal{M}_{\bar{\theta}}$ is reached along the shortest path (Fig. 3) [10], [11].

XIX. NUMERICAL SIMULATION

Fig. 4 shows a decimated sequence of synthetic images of a wire-framed cube with a significant amount of anisotropic and inhomogeneous Gaussian noise. On each grid point is centered the uncertainty ellipse indicating $\sqrt{2}$ times the standard devi-

ation in each orientation. Here, the weak perspective camera model (4) is assumed. The true object shape is shown bottom left; on bottom right is an optimal 3-D reconstruction with uncertainty ellipsoids indicating twice the standard deviation in each orientation (see [12], [13] for the details of the computation).

Fig. 5(a) shows the predicted normal form of the covariance matrix of all the parameters (the x, y, and z components of all the shape vectors $\{s_\alpha\}$, the motion parameters $\{t_\kappa, R_\kappa\}$, and the scales $\{\mu_\kappa\}$); the value of each element is displayed by its darkness. Fig. 5(b) is the description for the standard gauge (see (21) and (22)) obtained by applying the transformation (66) to Fig. 5(a). We see that these two descriptions are very different, yet the underlying uncertainty is the same. Fig. 5(c) shows the empirical covariance matrix computed by a random noise Monte Carlo simulation (400 runs) using the standard gauge. This demonstrates that our theoretical prediction describes the uncertainty very well.

XX. REAL IMAGE EXAMPLE

Fig. 6(a), (b) shows two real images. Here, we assume the perspective camera model (5) with unknown focal lengths. Selecting feature points marked in the images by hand, we reconstructed the 3-D structure by the method described in [8], [9]. Fig. 6(c) is a side view of the reconstructed points; wire-frames consisting of some of them are shown. On each reconstructed point is centered the uncertainty ellipsoid defined by the covariance matrix. All ellipsoids look like thin needles, indicating that the uncertainty is very large along the depth orientation.

This description is deceptive, however. This description is based on the standard gauge (21): the world coordinate system coincides with the first camera frame and the camera translation is normalized to unit length. This gauge hides the fact that the uncertainty is mostly due to the camera translation, not the object shape.

For example, if we take the centroid of the polyhedral object as the coordinate origin and normalize the root-mean-square distance to the vertices from the centroid to unit length, we obtain the description shown in Fig. 7(a). By construction, the uncertainty is almost symmetric with respect to the centroid. This description is better than Fig. 6(c) for seeing that the object shape has very little uncertainty.
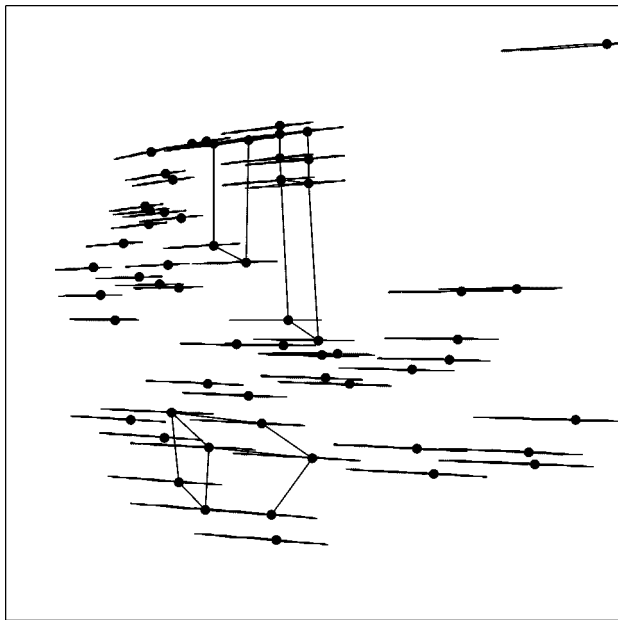
Fig. 7(b) is the uncertainty description for yet another gauge: one of the object vertex is taken to be the coordinate origin,
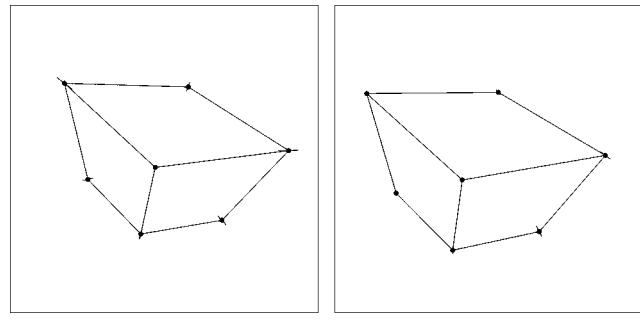
(a)

(b)

(c)

Fig. 6. (a), (b) Real images of an indoor scene. (c) 3-D reconstruction with uncertainty ellipsoids.

(a)　　　　　　　　　　(b)

Fig. 7. Gauge dependence of uncertainty description. (a) Normalization based on the centroid. (b) Normalization based on three vertices.

TABLE I
RELIABILITY OF GAUGE INVARIANTS

	ratio	angle (deg)
computed value	1.014	96.5
true value	1.000	90.0
predicted standard deviation	0.003	2.4

XXI. CONCLUDING REMARKS

We have presented a consistent theory for describing indeterminacy and uncertainty of 3-D reconstruction from a sequence of images. We have given a group-theoretical analysis of gauge transformations and gauges and extended the Cramer–Rao lower bound to incorporate internal indeterminacy. We also introduced the free-gauge approach and defined the normal form of a covariance matrix that is independent of particular gauges. Finally, we showed simulated and real-image examples to illustrate the effect of gauge freedom on uncertainty description.

REFERENCES

[1] H. Ballard and C. M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1982.
[2] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*. Cambridge, MA: MIT Press, 1993.
[3] W. Förstner, "Reliability analysis of parameter estimation in linear models with applications to mensuration problems in computer vision," *Comput. Vision Graphics Image Processing*, vol. 40, pp. 273–310, 1987.
[4] K. Kanatani, *Group-Theoretical Methods in Image Understanding*. Berlin, Germany: Springer-Verlag, 1990.
[5] ——, *Geometric Computation for Machine Vision*. Oxford, U.K.: Oxford Univ. Press, 1993.
[6] ——, *Statistical Optimization for Geometric Computation:Theory and Practice*. Amsterdam, The Netherlands: Elsevier, 1996.
[7] ——, "Cramer–Rao lower bounds for curve fitting," *Graphical Models Image Processing*, vol. 60, no. 2, pp. 93–99, 1998.
[8] ——, "Gauge-based reliability analysis of 3-D reconstruction from two uncalibrated perspective views," in *Proc. 15th Int. Conf. Pattern Recognition*, vol. 1, Barcelona, Spain, Sept. 2000, pp. 76–79.
[9] K. Kanatani and C. Matsunaga, "Closed-form expression for focal lengths from the fundamental matrix," in *Proc. 4th Asian Conf. Computer Vision*, vol. 1, Taipei, Taiwan, Jan. 2000, pp. 128–133.

another is taken to be $(1, 1, 0)$, and a third one is on the XY plane. By definition, the two points have no uncertainty.

These examples visualize the fact that uncertainty of individual quantities has no absolute meaning. Absolute meaning can be given only to gauge invariants. Typical gauge invariants are ratios of lengths and angles of lines. The ratio of two sides of the polyhedral objects and the angle they make are listed in Table I along with their true values and rough estimates of their theoretical standard deviations. Only by evaluating the uncertainty of gauge invariants in this way can we describe the degree of uncertainty in absolute terms.

[10] P. F. McLauchlan, "Gauge invariance in projective 3D reconstruction," in *IEEE Workshop on Multi-View Modeling and Analysis of Visual Scenes*, Fort Collins, CO, June 1999.

[11] ——, "Gauge independence in optimization algorithms for 3D vision," in *Vision Algorithms: Theory and Practice*, B. Triggs, A. Zisserman, and R. Szeliski, Eds. Berlin, Germany: Springer-Verlag, 2000, pp. 183–199.

[12] D. D. Morris and T. Kanade, "A unified factorization algorithm for points, line segments and planes with uncertainty models," in *Proc. 6th Int. Conf. Computer Vision*, Bombay, India, Jan. 1998, pp. 696–702.

[13] D. D. Morris, K. Kanatani, and T. Kanade, "Uncertainty modeling for optimal structure from motion," in *Vision Algorithms: Theory and Practice*, B. Triggs, A. Zisserman, and R. Szeliski, Eds. Berlin, Germany: Springer-Verlag, 2000, pp. 200–217.

[14] C. J. Poelman and T. Kanade, "A paraperspective factorization method for shape and motion recovery," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 206–218, Mar. 1997.

[15] M. Pollefeys, R. Koch, and L. Van Gool, "Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters," *Int. J. Comput. Vision*, vol. 32, no. 1, pp. 7–26, 1999.

[16] T. Poston and I. Stewart, *Catastrophe Theory and Its Applications*. London, U.K.: Pitman, 1978.

[17] J. Shi and C. Tomasi, "Good features to track," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Seattle, WA, June 1994, pp. 593–600.

[18] A. Singh, "An estimation-theoretic framework for image-flow computation," in *Proc. 3rd Int. Conf. Computer Vision*, Osaka, Japan, Dec. 1990, pp. 168–177.

[19] R. Szeliski and S. B. Kang, "Shape ambiguities in structure from motion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 506–512, May 1997.

[20] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography—A factorization method," *Int. J. Comput. Vision*, vol. 9, no. 2, pp. 137–154, 1992.

[21] B. Triggs, "Optimal estimation of matching constraint," in *3D Structure from Multiple Images of Large-Scale Environments (Lecture Notes in Computer Science)*, R. Koch and L. Van Gool, Eds. Berlin, Germany: Springer-Verlag, 1998, vol. 1506, pp. 63–77.

[22] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment—A modern synthesis," in *Vision Algorithms: Theory and Practice*, B. Triggs, A. Zisserman, and R. Szeliski, Eds. Berlin, Germany: Springer-Verlag, 2000, pp. 298–375.

[23] J. Weng, N. Ahuja, and T. S. Huang, "Optimal motion and structure estimation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 864–884, Sept. 1993.