To appear in CVPR 2001

Gauge Fixing for Accurate 3D Estimation

Daniel D. Morris
Northrop Grumman Corp.
1501 Ardmore Blvd.
Pittsburgh, PA 15221

Kenichi Kanatani
Dept. of Information Technology
Okayama University
Okayama 700-8530, Japan

Takeo Kanade
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

Computer vision techniques can estimate 3D shape from images, but usually only up to a scale factor. The scale factor must be obtained by a physical measurement of the scene or the camera motion. Using gauge theory, we show that how this scale factor is determined can significantly affect the accuracy of the estimated shape. And yet these considerations have been ignored in previous works where 3D shape accuracy is optimized. We investigate how scale fixing influences the accuracy of 3D reconstruction and determine what measurement should be made to maximize the shape accuracy.

1 Introduction

Our goal is to understand factors determining the uncertainties in 3D estimation from a set of cameras with unknown positions. In particular we want to know how fixing scale affects the uncertainty. Through understanding this we will obtain criteria for optimal scale fixing.

When 3D points are triangulated from known camera positions as in stereo, the uncertainties of the estimated points has been studied. The rule of thumb for points distant from the cameras is that the uncertainty is greater along the viewing direction than perpendicular to it, as illustrated in Fig. 1.

However, we are interested in the case where the camera positions and orientations are not initially known, but rather are estimated along with the 3D features. Here the same qualitative uncertainty analysis no longer applies for two reasons. First, how the coordinate system and scale are chosen can significantly affect the uncertainties, as shown in Fig. 2. Second, strong correlation between 3D feature points can be introduced when the the camera motions are also estimated. Consequently it is difficult to obtain a qualitative understanding of the uncertainty.

To simplify our analysis we will only consider uncertainties of lengths, which are invariant to choice of coordinate system, but depend on scale. We will seek to understand how uncertainty depends on both physical geometry, and on how the scale is fixed.

In applications of 3D estimation we often need to know the scale of the object, and hence must introduce at least one physical measurement into the estimation process. This could involve either measuring a length on the object or equivalently placing a measuring rod in the scene. In order to decide where to place the measuring rod we could analyze the results after 3D reconstruction is done, but then it is too late. Hence one of the goals of this paper is to use the understanding we develop to determine where a rod should be placed to maximize accuracy before doing 3D reconstruction. We call this the scale fixing problem.

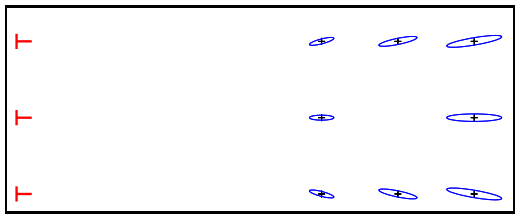


Figure 1: The scaled uncertainties of points estimated from three known camera positions and orientations. The ellipses represent the uncertainty, and have the characteristic elongation along the viewing direction.

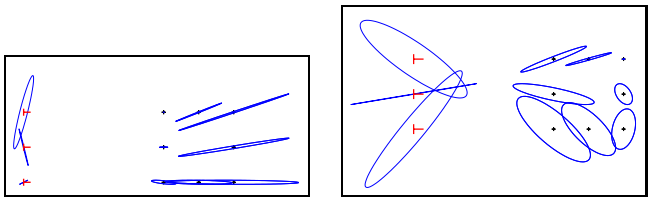


Figure 2: The scaled uncertainties of point and camera positions estimated with three cameras initially having unknown positions and orientations. The dramatic changes in uncertainties are due solely to the way the scale and coordinate system are chosen.

Previous work that dealt with shape uncertainties ignored the unknown scale factor [10, 12, 16] or stated that how it is determined is unimportant [14]. While past work has found optimal shape up to a scale factor, we argue that this is not good enough. In addition we need an optimal scale fixing method.

Our work uses recently developed gauge theory for handling indeterminacies in shape estimation [6, 8, 11, 15], and follows on from the older work in geodesy [2, 13] on free network design. We apply this to the scale fixing problem in computer vision.

2 Summary of Gauge Theory

First we summarize gauge theory for modeling indeterminacies in estimation problems; see [5, 6, 9, 11].

2.1 Gauge Orbits

In some estimation problems there are inherent indeterminacies in the parameters. In our case scale is one such indeterminacy, since the size of the object and positions of the cameras can all be rescaled to any value without changing the measurement. If the parameters are denoted with a point, $\boldsymbol{s}$, in a large dimensional space, $\mathcal{S}$, then we represent the indeterminacies as orbits filling this space. We call these *gauge orbits*, denoted $\mathcal{G}_{\boldsymbol{s}}$. All points on a gauge orbit are equivalent under our measurements[1]. A *gauge transformation*, g, takes one point, $\boldsymbol{s}' \in \mathcal{G}_{\boldsymbol{s}}$ to another point on the orbit: $\boldsymbol{s} \in \mathcal{G}_{\boldsymbol{s}}$ and is expressed as:

$$\boldsymbol{s} = g(\boldsymbol{s}'). \tag{1}$$

In our case scale acts as a linear gauge transformation, $g(\boldsymbol{s}') = a\boldsymbol{s}'$, and points on the gauge orbit correspond to models that are identical except by a scale factor, a:

$$\boldsymbol{s} = a\boldsymbol{s}' \,. \tag{2}$$

2.2 Constraints and Covariance Subspace

The covariance matrix provides a first order perturbation analysis around a point in parameter space. But when our solution is a whole gauge orbit, we need to constrain it to a unique point to do covariance analysis, as shown in Fig. 3. In our case, the gauge orbit is one dimensional, and so a single constraint on the shape will suffice, which we denote as $c(\boldsymbol{s}) = 0$.

This constraint defines a *gauge manifold*, or simply *gauge*, $\mathcal{C}$, of points satisfying it. Enforcing a constraint like this, we call *gauge fixing* or *choosing a gauge*. Perturbations must be in the tangent space to this manifold.

2.3 Geometric Equivalence

One of the key results derived in gauge theory is that given a perturbation, $\Delta\boldsymbol{s}$, at a point on the gauge orbit, then incrementing it with any perturbation in the tangent space to the gauge orbit, $T[\mathcal{G}_{\boldsymbol{s}}]$, does not alter its geometric meaning with respect to the measurements. This is an extension of the idea that perturbing a point on a gauge orbit does not change its measurement. This is captured in the following theorem which is proved in [6, 9].

Theorem 2.1. *Two perturbations, $\Delta\boldsymbol{s} \in T[\mathcal{C}]$ and $\Delta\boldsymbol{s}' \in T[\mathcal{C}']$, where $\boldsymbol{s} = a\boldsymbol{s}'$, are geometrically equivalent if and only if:*

$$\boldsymbol{u}^\top(\Delta\boldsymbol{s} - a\Delta\boldsymbol{s}') = 0 \quad \forall \boldsymbol{u} \in T[\mathcal{G}_{\boldsymbol{s}}]^\perp \,. \tag{3}$$

Geometric equivalence is illustrated in Fig. 4, and it requires that the difference vector, $\Delta\boldsymbol{s} - a\Delta\boldsymbol{s}'$, is in the tangent space, $T[\mathcal{G}_{\boldsymbol{s}}]$. Thus we refer to this tangent space as consisting of all the *gauge freedoms* of the point.

[1] Mathematicians call such an orbit a *leaf*, and a space filled with leafs a *foliation*.

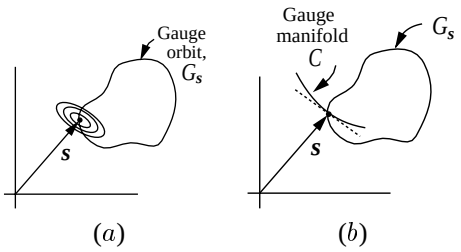


Figure 3: Plot (a) shows the contradiction we obtain when we use a full rank covariance to describe perturbations of a point on a gauge orbit. Ellipses of different probability intersect the gauge orbit, all of whose points are equivalent and hence have equal probability. Instead, we must impose a constraint that reduces our parameter space to the gauge manifold, $\mathcal{C}$, which intersects the gauge orbit at a single point, as shown in (b). Perturbations are now restricted to the tangent plane to $\mathcal{C}$.

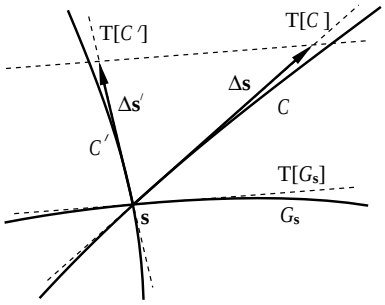


Figure 4: Perturbations in the tangent planes to two gauges, $\mathcal{C}$ and $\mathcal{C}'$, are shown when $a = 1$. These perturbations are geometrically equivalent when their difference, $\Delta\boldsymbol{s} - a\Delta\boldsymbol{s}'$, is orthogonal to all vectors $\boldsymbol{u} \in T[\mathcal{G}_{\boldsymbol{s}}]^\perp$.

3 Transforming Uncertainties

We will assume our problem is the following. Optimal 3D estimation has been done for a set of points and a set of camera positions and orientations, but the scale, rotation and translation are arbitrarily set. We will only consider distances between feature points which are invariant to rotation and translation, and henceforth can ignore rotation and translation gauge freedoms. We would like to find the scale in a way that maximizes the accuracy of our shape estimation.

For our purposes, we will only need to represent shape parameters, even though to obtain these we initially had to estimate camera motion. Let us represent the shape of our object with a vector, $\boldsymbol{s}' = (\boldsymbol{s}'_1, \ldots, \boldsymbol{s}'_N)$, that contains the 3D coordinates of the N individual feature points $\boldsymbol{s}'_i$. The covariance is obtained in the standard way by inverting the Fisher information matrix, which is the Hessian of the appropriate cost function for optimal shape and motion estimation. If $\boldsymbol{m}$ represents the camera motion parameters, we obtain the shape covariance, $\boldsymbol{V}_{\boldsymbol{s}'}$, from:

$$\begin{pmatrix} \boldsymbol{V}_{\boldsymbol{s}'} & \boldsymbol{V}_{\boldsymbol{m}'\boldsymbol{s}'} \\ \boldsymbol{V}_{\boldsymbol{s}'\boldsymbol{m}'} & \boldsymbol{V}_{\boldsymbol{m}'} \end{pmatrix} = \begin{pmatrix} H_{\boldsymbol{s}} & H_{\boldsymbol{s}\boldsymbol{m}} \\ H_{\boldsymbol{m}\boldsymbol{s}} & H_{\boldsymbol{m}} \end{pmatrix}^{-} \tag{4}$$

with the shape and motion components of the Hessian on the right, and $(\cdot)^-$ indicating the generalized inverse[2]. In general $\boldsymbol{V}_{\boldsymbol{s}'}$ may contain strong correlation terms between the shape parameters.

To find the scale we will measure the distance, d, between two feature points, i and j, on the real object. This gives us a constraint:

$$c(s) = d - \|\boldsymbol{s}_i - \boldsymbol{s}_j\| = 0\,, \tag{5}$$

which defines our gauge manifold $\mathcal{C}$. The appropriate rescaling factor is then: $a = d/\|\boldsymbol{s}'_i - \boldsymbol{s}'_j\|$. Once we obtain scale, we can transform $\boldsymbol{V}_{\boldsymbol{s}'}$ to the tangent plane of the new gauge, $\mathcal{C}$, and so obtain our transformed covariance $\boldsymbol{V}_{\boldsymbol{s}}$ as described in the following section.

3.1 Gauge Fixing

We can now derive the transformation of our original perturbations, $\Delta\boldsymbol{s}'$, described by $\boldsymbol{V}_{\boldsymbol{s}'}$, into the new gauge constraint $\mathcal{C}$, such that geometric equivalence is maintained. To do this we follow the gauge theory method.

One solution to the geometric equivalence relationship is a simple rescaling of the perturbation $\Delta\boldsymbol{s} = a\Delta\boldsymbol{s}'$. But in general the resulting perturbation will not be tangent to the new gauge, $\mathcal{C}$. The full solution is:

$$\Delta\boldsymbol{s} = a\Delta\boldsymbol{s}' + \boldsymbol{b} \tag{6}$$

where $\boldsymbol{b}$ is a vector in the tangent to the gauge orbit: $T[\mathcal{G}_{\boldsymbol{s}}]$. The tangent is given by $\partial\boldsymbol{s}/\partial a = \boldsymbol{s}'$, and hence we can write $\boldsymbol{b} = x\boldsymbol{s}$ with some unknown coefficient x.

We want to choose x to satisfy the gauge constraint, so we let $\boldsymbol{v}$ be a vector orthogonal to the gauge tangent space, $\boldsymbol{v} \in T[\mathcal{C}]^{\perp}$. One such vector is given by the gradient of the constraint $\boldsymbol{v} = \nabla_{\boldsymbol{s}} c(\boldsymbol{s})$, where $c(\boldsymbol{s}) = 0$ is defined by Eq. (5). Since $\boldsymbol{v}$ is orthogonal to $\Delta\boldsymbol{s}$, our gauge constraint is satisfied if:

$$\boldsymbol{v}^\top \Delta\boldsymbol{s} = 0\,. \tag{7}$$

Applying this to Eq. (6), letting $\boldsymbol{b} = x\boldsymbol{s}$, and solving for x, we obtain: $x = -(\boldsymbol{v}^\top\boldsymbol{s}')^{-1}\boldsymbol{v}^\top a\Delta\boldsymbol{s}'$. Then substituting this into Eq. (6) we get:

$$\Delta\boldsymbol{s} = a\Delta\boldsymbol{s}' - \boldsymbol{s}'(\boldsymbol{v}^\top\boldsymbol{s}')^{-1}\boldsymbol{v}^\top a\Delta\boldsymbol{s}' = a\boldsymbol{Q}\Delta\boldsymbol{s}'\,, \tag{8}$$

where

$$\boldsymbol{Q} = \boldsymbol{I} - \frac{\boldsymbol{s}'\boldsymbol{v}^\top}{\boldsymbol{v}^\top\boldsymbol{s}'}\,. \tag{9}$$

The matrix $\boldsymbol{Q}$ is an oblique projection operator[3], and it corresponds to Baarda's S-Transformation [2]. It takes any scaled perturbation $a\Delta\boldsymbol{s}'$ along the tangent plane to a geometrically equivalent perturbation $\Delta\boldsymbol{s}$ in a new gauge $\mathcal{C}$.

[2] If this is not full rank we typically take the Moore-Penrose inverse. See [9] for details.

[3] It is because $\boldsymbol{Q}$ has this form that we can drop the motion elements, $\boldsymbol{V}_{\boldsymbol{s}'\boldsymbol{m}'}$ etc., from the full covariance. Only the elements that interact with the constraint vector, $\boldsymbol{v}$, and that we want to estimate need to be included, as one can easily confirm by expanding the corresponding full expression.

We can then transform our covariance to $T[\mathcal{C}]$:

$$\boldsymbol{V}_{\boldsymbol{s}} = E[\Delta\boldsymbol{s}\,\Delta\boldsymbol{s}^\top] = a^2\boldsymbol{Q}\boldsymbol{V}_{\boldsymbol{s}'}\boldsymbol{Q}^\top\,, \tag{10}$$

where $\boldsymbol{V}_{\boldsymbol{s}}$ gives the uncertainty of the shape after we make measurement d.

3.2 Measurement Uncertainty

If there is uncertainty in our measurement d, we can modify Eq. (2) to incorporate this:

$$\Delta\boldsymbol{s} = a\boldsymbol{Q}\boldsymbol{s}' + \boldsymbol{s}'\Delta a\,. \tag{11}$$

The scale is $a = d/d'$ where $d' = \|\boldsymbol{s}'_i - \boldsymbol{s}'_j\|$, and so $\Delta a = \Delta d/d'$. If d is measured with standard deviation σ_m, then we obtain:

$$\boldsymbol{V}_{\boldsymbol{s}} = a^2\boldsymbol{Q}\boldsymbol{V}_{\boldsymbol{s}'}\boldsymbol{Q}^\top + \sigma_m^2\frac{\boldsymbol{s}'\boldsymbol{s}'^\top}{d'^2}\,. \tag{12}$$

We see that this measurement error has less effect the larger length d' is compared to the object.

4 Finding the Best Gauge Constraint

We want to find the best gauge constraint. Our criterion is that we want to estimate a certain length on the object, $e = \|\boldsymbol{s}_k - \boldsymbol{s}_l\|$, with the greatest accuracy, and that we can set the scale factor by measuring another length, d. What qualities should d have to minimize the variance of e?

Let $\boldsymbol{g}' = (e' \;\; d')^\top$ and $\boldsymbol{g} = a\boldsymbol{g}'$, and so we can write:

$$\boldsymbol{V}_{\boldsymbol{g}'} = \nabla_{\boldsymbol{s}'}\boldsymbol{g}'^\top\boldsymbol{V}_{\boldsymbol{s}'}\nabla_{\boldsymbol{s}'}\boldsymbol{g}' \equiv \begin{pmatrix} \sigma_{e'}^2 & \sigma_{e'd'} \\ \sigma_{d'e'} & \sigma_{d'}^2 \end{pmatrix}. \tag{13}$$

If we measure d with variance σ_m^2, we obtain from Eq. (12):

$$\boldsymbol{V}_{\boldsymbol{g}} = a^2\boldsymbol{Q}_{\boldsymbol{g}}\boldsymbol{V}_{\boldsymbol{g}'}\boldsymbol{Q}_{\boldsymbol{g}}{}^\top + \sigma_m^2\frac{\boldsymbol{g}\boldsymbol{g}^\top}{d^2}. \tag{14}$$

Here our scale factor gives us a gauge freedom: $\partial\boldsymbol{g}/\partial a = \boldsymbol{g}'$, and our measurement d gives us a local constraint: $\boldsymbol{v} = (0 \;\; 1)^\top$, and so

$$\boldsymbol{Q}_{\boldsymbol{g}} = \begin{pmatrix} 1 & -e/d \\ 0 & 0 \end{pmatrix}. \tag{15}$$

We want to minimize the variance of e, which can be obtained algebraically from Eq. (14) as:

$$\sigma_e^2 = a^2(\sigma_{e'}{}^2 - 2\frac{e}{d}\sigma_{e'd'} + \left(\frac{e}{d}\right)^2\sigma_{d'}^2) + \left(\frac{e}{d}\right)^2\sigma_m^2\,. \tag{16}$$

This is quadratic in the ratio e/d, and its minimum has two cases. The first is when the noise is uncorrelated or anti-correlated, ($\sigma_{e'd'} \leq 0$). The variance, σ_e^2, is reduced when the ratio, e/d, is reduced. Thus the longer the length we measure, d, the more accurate our estimate for e is. The second case is when the noise is positively correlated, $\sigma_{e'd'} > 0$. The ratio that minimizes σ_e^2 is then:

$$\frac{e}{d} = \frac{a^2\sigma_{e'd'}}{a^2\sigma_{d'}^2 + \sigma_m^2}\,. \tag{17}$$

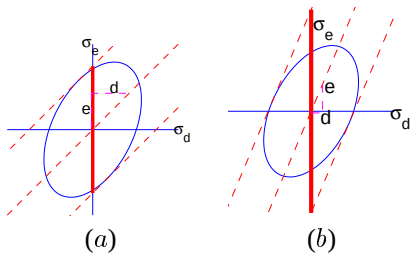


Figure 5: Illustrated are the projections of a covariance $\boldsymbol{V}_{\boldsymbol{g}'}$ onto a single axis to give us σ_e, as described in Eq. (14) with σ_m=0. In (a) the gauge freedom direction is given by e/d = 1, whereas in (b) $e/d > 1$. Notice that the gauge freedom direction, e/d, significantly affects the magnitude of the projection onto the e axis.

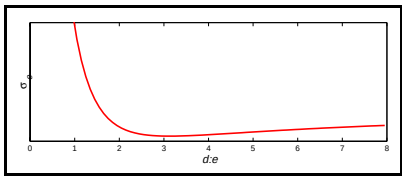


Figure 6: The plot shows the standard deviation, σ_e, of the predicted line length, e, as a function of the ratio of the measured line length to the predicted line length, $d : e$. In general the larger d the better the estimate, except beyond a certain value after which the uncertainty increases moderately to an asymptote.

If the noise is perfectly correlated, namely $\sigma_{e'd'} = \sigma_{e'}\sigma_{d'}$, and σ_m^2 = 0, and the ratio, e/d, is given by Eq. (17), then the length e will be perfectly estimated.

We can also understand Eq. (14) geometrically using our oblique projection interpretation. Let us represent $\boldsymbol{V}_{\boldsymbol{g}'}$ with a 2D ellipse giving the standard deviation of d' and e'. The oblique projection operator, $\boldsymbol{Q}_{\boldsymbol{g}}$, projects this along the gauge freedom onto the e axis. The gauge freedom is a line whose slope is e/d as shown in Fig. 5. The steeper the slope, or the smaller d is compared to e, the larger the resulting uncertainty, σ_e, is. The relationship between ratio $d : e$ and σ_e is plotted for this example in Fig. 6.

A further interesting consequence of the projection shown in Fig. 5(a) is that if the ratio, e/d = 1, then the gauge freedom will be at 45°, and the projection of the ellipse onto the coordinate axes will be the same if either d or e is fixed. This means that if the lines are the same length, then holding one line fixed and predicting the other, or holding the other fixed and predicting the first will give exactly the same uncertainty. This symmetry property, depending only on the length, will simplify gauge fixing calculations for lines of equal length.

4.1 Line Orientation in Sphere Estimation

It is not only line length that affects the accuracy of gauge fixing. In this section we consider relative orientation of lines. For simplicity we take $\sigma_m = 0$.

A synthetic sequence of a sphere is created by orthogonally intersecting 3 circles. Lines are formed by connecting antipodal points, and are all of the same length. The camera faces the sphere and five images are taken, as illustrated in Fig. 7. Uniform, zero-mean, uncorrelated Gaussian noise is assumed. Our goal is to find which lines provide the best gauge fixing constraints. To do this we will test the four diameters labeled 1, 7 and 18 in the figure.

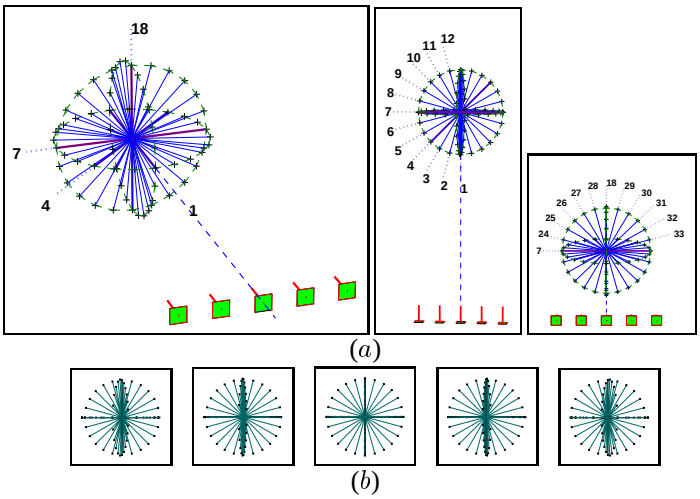


Figure 7: (a) Views of a sphere created by intersecting 3 circles orthogonally, and five cameras. Diameter 1 is along the viewing direction and 7 and 18 are in the plane orthogonal to the viewing direction. The dashed circles are only to aid in viewing. (b) Images seen by the cameras, with lines connecting points added afterwards.

As a first step in fixing line 1, we calculate $\boldsymbol{V}_{\boldsymbol{g}'}$ between line 1 and each of the other lines, with d' representing line 1, and e' in turn representing each of the other lines. The uncertainty of each line, $\sigma_{e'}$ is shown in Fig. 8(a) along with the relative correlation between it and line 1. We rescale and obliquely project it according to Eq. (16) to get σ_e, the standard deviation of each line given line 1, which shown in Fig. 8(b). Figure 9 shows the result for lines 7 and 18.

We see from Fig. 8(a) that lines along the viewing direction, (numbers 2, 3, 11, 12, 13, 22, 23), have the greatest initial uncertainty when calculated from $\boldsymbol{V}_{\boldsymbol{s}'}$[4]. However, they are also very strongly correlated to line 1. Hence if we compare their uncertainties after we measure line 1, (Fig. 8(b)), to their uncertainties after we measure line 7, (Fig. 9), we find that they are estimated to almost twice the accuracy when line 1 is measured.

Again comparing Fig. 8(b) and Fig. 9, we see that lines orthogonal to the viewing direction are much more accurately predicted when line 7, or any of the lines orthogonal to the viewing direction, is measured. The orientation in this plane makes no difference to the accuracy.

[4]The values in Fig. 8(a) could vary depending on how we calculate $\boldsymbol{V}_{\boldsymbol{s}'}$, but they still reveal the structure of the uncertainty.

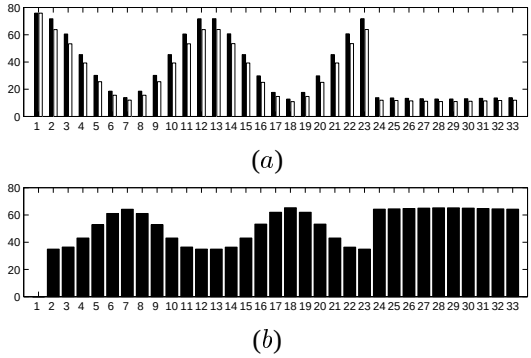


Figure 8: The black bars in (a) show $\sigma_{e'}$, the uncertainty of each line calculated from $\boldsymbol{V}_{\boldsymbol{s}'}$. The white bars give the relative correlation, $\sigma_{e'd'}/\sigma_{d'}$, between each line and line 1. (b) Rescaling and performing our oblique projection in Eq. (16) for each case, we obtain the uncertainty, σ_e, for each measured line.

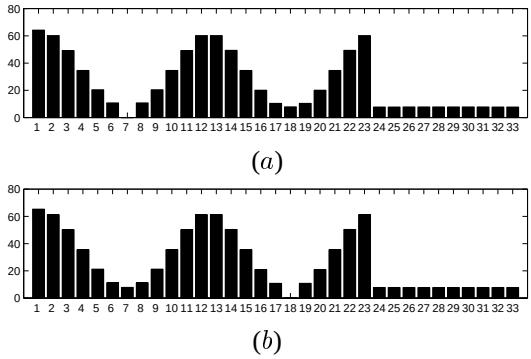


Figure 9: The uncertainties, σ_e, in predicting each line when (a) line 7, and (b) line 18 is measured. The plot for line 4 is intermediate between this and that in Fig. 8.

4.2 Reducing Correlation Effects

In the previous example our sphere is the whole object. We would like to know how our results change when the lines we are interested in are only part of the object that we are reconstructing. Experiments show that adding more features on or close to the sphere will not qualitatively change the results [9]. But if our object is large compared to the sphere, the results do change. To show this we added one hundred randomly positioned features, within a rectangular region, that are assumed to be rigidly attached to the sphere, as shown in Fig. 10(a).

We find that when line 1 is measured, the uncertainty pattern is inverted: compare Figs. 8(b) and 10(c). Now lines in the plane perpendicular to the viewing direction are more accurately predicted than lines close to line 1.

The immediate reason for the inversion of this uncertainty pattern can be found in Figs. 8(a) and 10(b). In the case of the single sphere, line 1 is strongly correlated with the other lines, whereas in the case where the sphere is a small part of the object, the correlation between line 1 and the other lines is virtually gone. (No correlation means we will get the same uncertainty pattern as the known camera case shown in Fig. 1, see Matthies and Shafer [7]).

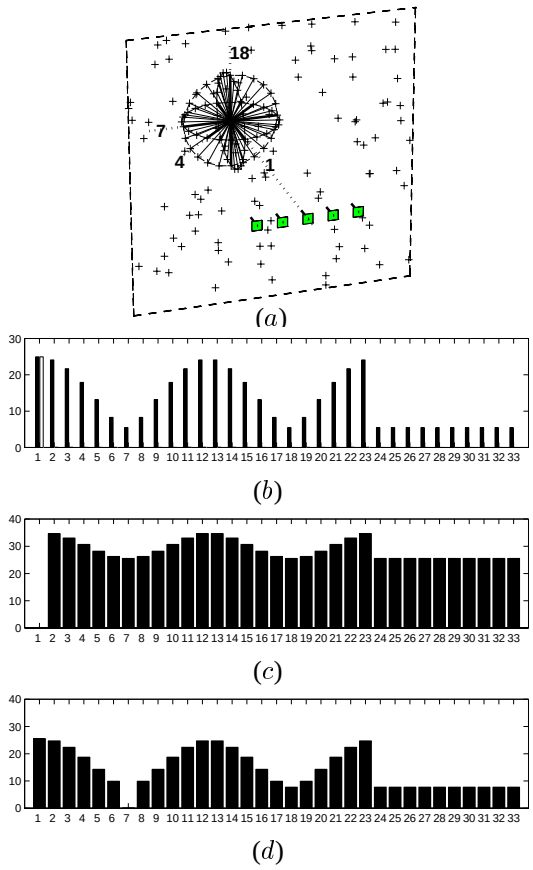
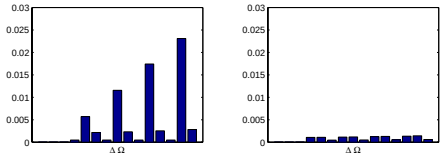


Figure 10: (a) A random set of 3D points in a large rectangular planar region are added to the 3D sphere. (b) Uncertainties from $\boldsymbol{V}_{\boldsymbol{g}'}$ are shown. Black lines give $\sigma_{e'}$ for each line and the white lines show the correlation with line 1. (c) The resulting uncertainty, σ_e, for each line when line 1 is measured, and (d) σ_e when line 7 is measured.

We would like to know why there is large correlation in one case and not in the other. This comes from the original calculation of the $\boldsymbol{V}_{\boldsymbol{s}'}$ covariance in Eq. (4). If the camera positions and orientations are known exactly, then $\boldsymbol{V}_{\boldsymbol{s}'} = H_{\boldsymbol{s}}^{-1}$, and it turns out that the points are all uncorrelated. However, when there is uncertainty in motion, this ends up, in part, as correlation terms in $\boldsymbol{V}_{\boldsymbol{s}'}$ in Eq. (4). When additional points are added as in Fig. 10(a), the greater field of view enables rotations to be estimated with much less uncertainty as shown in Fig. 11, and so gives less correlation in $\boldsymbol{V}_{\boldsymbol{s}'}$.

We conclude from this set of experiments that when our object is large in width and height, compared to the set of lines we want to estimate, then it is best to measure lines orthogonal to the viewing direction. On the other hand, for objects with lines all clustered around a single point, lines along the viewing direction are best predicted by measur-



(a) (b)

Figure 11: The uncertainty in rotation for the single sphere from Fig. 7 is shown in (a), and for the sphere with additional points from Fig. 10 in (b).

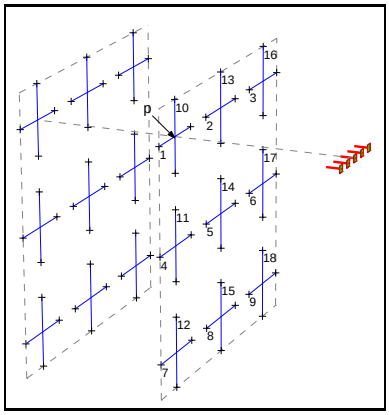


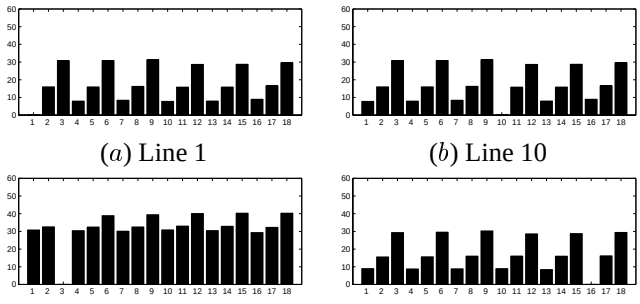
Figure 12: Eighteen lines of equal length in a plane are labeled.

ing other lines along the viewing direction. This is illustrated in Table 1.

4.3 Position in the Plane

Next we would like to find out how position of a line in a plane affects its usefulness for gauge fixing. A synthetic object made from two planes is shown in Fig. 12.

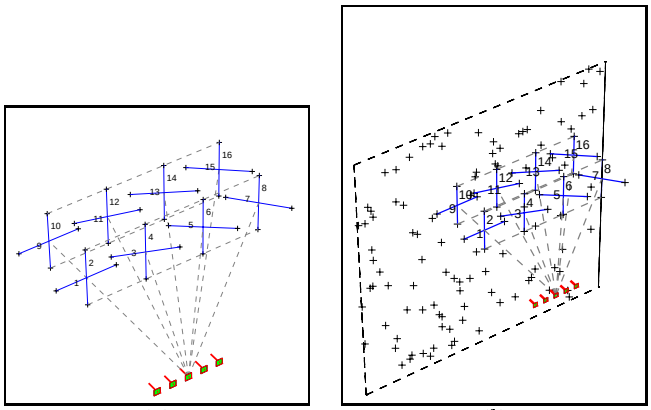
Figure 13 shows the resulting uncertainties when a selection of the lines is used to fix the scale. We notice first that irrespective of which line is used to fix the scale, the accuracy order of the lines stays the same (of course excluding the line that is fixed in each case). That is, lines 1, 4, 7, 10, 13 and 16 are always most accurately predicted, lines 2, 5, 8, 11, 14 and 17 are next, and lines 3, 6, 9, 12, 15 and 18 are least accurately predicted. This result applies



(a) Line 1 (b) Line 10

(c) Line 3 (d) Line 16

Figure 13: The uncertainty of line prediction is shown for each of the labeled lines.



(a) (b)

Figure 14: Five views of a shape made from two planes. In both plots the mid-points of lines 1 through 8 are on one plane and the rest on the plane behind this. Lines 2, 4, 6 and 8 are fully on the front plane, and lines 1, 3, 5 and 7 are of the same length but oriented so that they are orthogonal to their viewing directions. In (b) an additional set of points are added to make the object large compared to the sphere.

in reverse too, that is, better predicted lines are better for gauge fixing.

One might suppose that the reason for this pattern is due to the projected line length in the images. But this cannot be the case since for a perspective camera all the lines in the plane have the same projected line length. Our next experiment will investigate reasons for this effect.

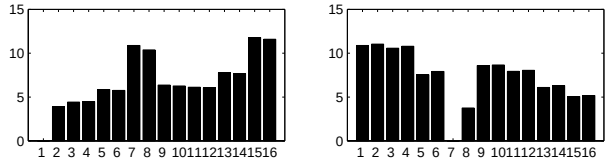
4.4 Viewing Directions and Perpendicular Lines

Consider the row consisting of lines 1, 2, 3, 10, 13, and 16 in Fig. 12. We find that fixing any one of the lines 10, 13 or 16, generates equal accuracy in predicting other lines, but that fixing one of lines 1, 2 or 3, gives increasingly poor accuracy. To understand this we created the following experiment. Here *viewing direction* of a line refers to the line from the center camera to the center of the line.

We took a row of lines from the parallel planes example, shown in Fig. 12, and, keeping their mid-points fixed, we swivelled the lines so that each one is orthogonal to its viewing direction. Thus we obtained the shape shown in Fig. 14(a). The results of measuring lines 1 and 7 are shown in Fig. 15. We see from this plot that lines that share a viewing direction, and are orthogonal to it, have approximately the same uncertainty. In this example the height of the object is small and so the line lengths are strongly correlated with each other.

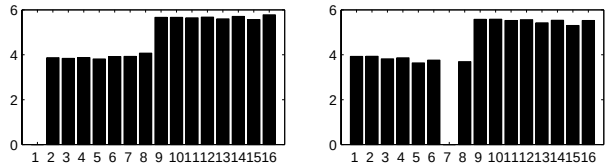
In order to remove line correlation, and so make our setup analogous to the previous parallel plane example, we added a set of random points to our shape and obtained the new object shown in Fig. 14(b). The resulting uncertainties are shown in Fig. 16. We see that all lines with mid-points on the front plane have the same uncertainty.

This allows us to explain our result in section 4.3. The



(a) (b)

Figure 15: The uncertainty of the line prediction for the shape in Fig. 14(a) is shown in (a) for line 1, and in (b) for line 7.



(a) (b)

Figure 16: The uncertainty of the line prediction for the shape in Fig. 14(b) is shown in (a) for line 1, and in (b) for line 7.

difference in pattern between Figs. 16 and 13 is due only to line orientation with respect to the viewing direction. Hence this plays a determining role in uncertainty. The lines in the plane that provided the best gauge fixing constraints were precisely those that are orthogonal to their viewing directions, and those that gave the greatest uncertainty were those having the greatest angle from the plane orthogonal to its viewing direction. This is the difference between lines 3 and 16 in Fig. 12. For a fuller set of experimental results see [9].

4.5 Desk Reconstruction

Our next experiment involved the reconstruction of a computer and desk. Seven images were taken of it using a digital camera and features were registered in the sequence as shown in Fig. 17. A 3D reconstruction was then performed and the result is shown in Fig. 18.

The first goal for this sequence was to estimate the width of the computer monitor labeled as line 1. We measured a number of lengths in the scene so see how each of these would do in fixing the scale for predicting line 1. All were assumed to have 0.5mm measurement uncertainty.

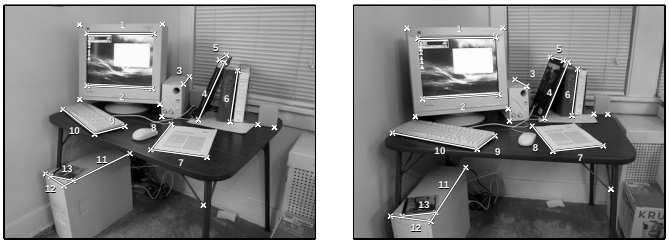


Figure 17: The two end-images from a seven image sequence of a computer and desk with lines marked.

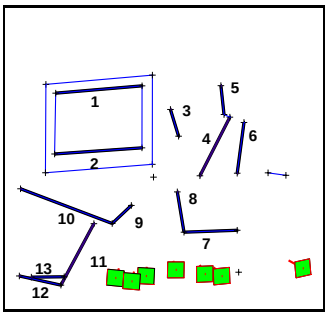


Figure 18: The 3D reconstruction of the desk is shown along with the camera positions.

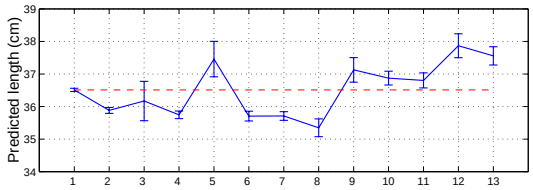


Figure 19: The predicted monitor width is shown, along with its uncertainty in standard deviations for each of the thirteen lines, calculated using Eq. (16).

The most promising lines for optimizing the accuracy of predicting line 1 through gauge fixing are 2, 4, 6 and 7 since they are all long and approximately orthogonal to their viewing directions. Lines 10 and 11 are even longer,and so better, although being oriented partly along the viewing direction this advantage may be somewhat offset. We expect 3 and 5 to be poor since they are oriented along their viewing directions and 9, 12 and 13 also to be poor since they are short. Figure 19 sustains all of these predictions.

Now instead imagine we want to know the length of the sub-woofer, indicated as line 3 in the image. The key question here is whether it would be better to know line 5, which is parallel to line 3, or else one of the lines orthogonal to the viewing direction. Since the desk has a large field of view in both height and width compared to the line lengths, we expect that the lines are relatively uncorrelated and hence that lines orthogonal to the viewing direction would be better than line 5. This is confirmed in Fig. 20.

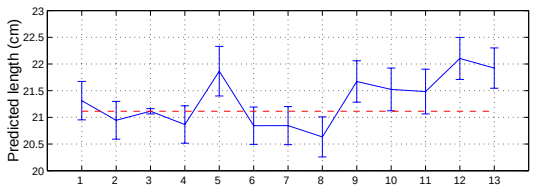


Figure 20: The predicted length of the sub-woofer is shown, along with its uncertainty in standard deviations for each of the thirteen lines, calculated using Eq. (16).

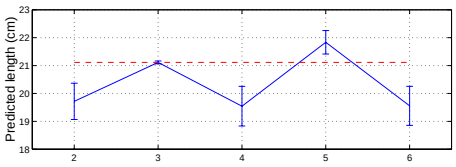


Figure 21: The predicted length of the sub-woofer is shown, along with its uncertainty in standard deviations for just lines 2, 3, 4, 5 and 6, calculated using Eq. (16).

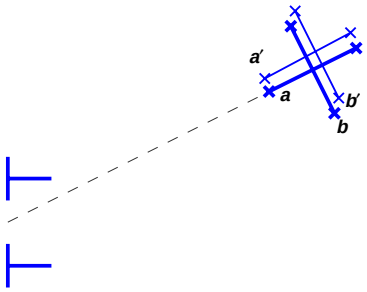


Figure 22: Four equal-length lines in an object, a, a', b, and b', with the first two being along the viewing direction, and the second two orthogonal to it. Table 1 shows the prediction accuracies.

But let us consider another situation in which the only features that we tracked are those belonging to lines 2, 3, 4, 5 and 6. We can do shape estimation with just these data, and then estimate the sub-woofer length by measuring one of the other lengths. With points from just these lines, the object has a relatively small field of view compared to the line lengths. Hence we might expect that measuring line 5 would give us the greatest accuracy. We see in Fig. 21 that this is indeed the case.

5 Conclusion

We have investigated how the covariances of object lengths are affected by fixing the scale. We found out how length, position and orientation of a line each affect the line's utility for fixing the scale. In doing this we distinguished objects having a large field of view from those having a small field of view relative to line length, as this determines the relative accuracy of camera motion estimation and hence the amount of correlation between the lines. Table 1 summarizes how line orientation can effect accuracy in scale fixing.

It is important to considering how the scale will be fixed when doing shape estimation. One way to fis scale is to place a measuring rod of known length in the scene prior to taking images, and include it as part of the object. The qualitative understanding obtained in this paper allows us to place the measuring rod in the scene at the position and orientation that will maximize the final 3D accuracy of the part of the object that we most want to estimate.

These results are also useful for optimally designing calibration objects for multi-camera systems or mobile robots. Future work will involve extending these principles to use multiple lines for scale fixing.

Table 1: The table shows which of line a or b is more accurately estimated when a' or b' is fixed in Fig. 22. As explained in section 4.2 this depends on the size of the complete object of which these lines are only a part.

	Small object (high correlation)	Large object (low correlation)
a' fixed	a	b
b' fixed	b	b

References

[1] K. B. Atkinson, *Close Range Photogrammetry and Machine Vision*, Whittles Publishing, Caithness, Scotland, (1996).

[2] W. Baarda, *S-Transformations and Criterion Matrices*, volume Band 5 der Reihe 1, Netherlands Geodetic Commission, (1973).

[3] R. M. Haralick, Propagating covariance in computer vision, *Int. J. of Patt. Recog. and AI.*, **10**(5), pp. 561–572, (1996).

[4] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier, Amsterdam, (1996).

[5] K. Kanatani & D. D. Morris, Gauges and gauge transformations in 3-D reconstruction from a sequence of images, in *Proc. Fourth Asian Conf. Comp. Vision*, volume 2, pp. 1046–1051, Taiwan, (2000).

[6] K. Kanatani & D. D. Morris, Gauge and gauge transformations for uncertainty description of geometric structure with indeterminacy, *IEEE Transactions on Information Theory*, **47**(5), pp. 2017–2028, (2001).

[7] L. Matties & S. A. Shafer, Error modeling in stereo navigation, *IEEE Journal of Robotics and Automation*, **RA-3**(3), pp. 239–48, (June 1987).

[8] P. F. McLauchlan, Gauge independence in optimization algorithms for 3D vision, in *Vision Algorithms: Theory and Practice*, pp. 183–199, Springer, Berlin, (2000).

[9] D. D. Morris, *Gauge Freedoms and Uncertainty Modeling for 3D Computer Vision*, Ph.D. thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, (March 2001), CMU-RI-TR-01-15.

[10] D. D. Morris & T. Kanade, A unified factorization algorithm for points, line segments and planes with uncertainty models, in *Proc. Sixth Int. Conf. Comp. Vision*, pp. 696–702, Bombay, India, (January 1998).

[11] D. D. Morris, K. Kanatani & T. Kanade, Uncertainty modeling for optimal structure from motion, in *Vision Algorithms: Theory and Practice*, pp. 200–217, Springer, Berlin, (2000).

[12] R. Szeliski & S. B. Kang, Shape ambiguities in structure from motion, *IEEE Trans. Patt. Anal. Mach. Intelligence*, **19**(5), pp. 506–512, (May 1997).

[13] P. Teunissen, Zero order design: Generalized inverses, adjustment, the datum problem and s-transformations, in *Optimization and Design of Geodetic Networks*, pp. 11–55, Springer-Verlag, Berlin, (1985), chap. 1 in book edited by E. W. Garafarend and F. Sanso.

[14] J. I. Thomas, A. Hanson & J. Oliensis, Refining 3D reconstructions: A theoretical and experimental study of the effect of cross-correlations, *CVGIP*, **60**(3), pp. 359–370, (November 1994).

[15] B. Triggs, P. F. McLauchlan, R. I. Hartley & A. W. Fitzibbon, Bundle adjustment – a modern synthesis, in *Vision Algorithms: Theory and Practice*, pp. 298–375, Springer, Berlin, (2000).

[16] J. Weng, N. Ahuja & T. Huang, Optimal motion and structure estimation, *IEEE Trans. Patt. Anal. Mach. Intelligence*, **15**(9), pp. 864–884, (1993).