3D Reconstruction Based on Homography Mapping*

Zhongfei Zhang†
CEDAR, State U. of NY
Buffalo, NY 14228-2567
zhongfei@cedar.buffalo.edu

Allen R. Hanson
Department of Computer Science
Univ. of Mass., Amherst, MA 01003
hanson@cs.umass.edu

Abstract

Previous work shows that based on the fundamental matrix from two views, 3D structures can be recovered up to an unknown projectivity. In this paper, we show that based on four ***coplanar*** *correspondences of two* ***externally*** *uncalibrated cameras, 3D reconstruction can be achieved in Euclidean space with only one uniform scale factor and up to two real solutions. It is shown that this scale factor is the physical distance from the camera center to the plane formed by the four points in 3D space. Consequently, if this distance is known a priori, then the 3D structure can be completely determined. In order to disambiguate the two solutions, a third view is required in general to give a unique solution. In practice, since the real data are always corrupted with noise, more coplanar correspondences are used and a least squares solution is applied to obtain the estimation of the homography matrix. Experimental results on both simulated and real data show that this reconstruction algorithm works reasonably robustly. It is also shown that this algorithm is optimal both in terms of the minimum number of required correspondences and in terms of the assumption made for the internal calibration.*

1 Introduction

3D reconstruction is an important and difficult problem in computer vision. Over the years, this problem has been related to structure from motion, stereo vision, pose determination, etc., and has been addressed by many researchers[BH87, HGC92, Kum92]. Owing to the wide spectrum of different constraints on camera geometry, the problem of 3D reconstruction can be formulated and solved in very different ways. On the one extreme, both internal calibration and external pose geometry are known. In this case, the problem of 3D reconstruction can be solved by simple triangulation techniques, such as Blostein and Huang[BH87]. On the other extreme, neither internal calibration nor external pose geometry is known. This is a problem of 3D reconstruction based on two *completely* uncalibrated images. Faugeras[Fau92] and Hartley et al[HGC92] are two of the earliest efforts to address the problem. They have shown that based on the fundamental matrix of two views, 3D structures can be recovered up to an unknown projectivity. More recent work in this direction can be found in Shashua[Sha93].

In this paper, we address the problem of 3D reconstruction by focusing on a "middle case" of the spectrum of possible camera geometry constraints. Specifically, we investigate the problem of 3D reconstruction based on *externally* uncalibrated cameras while assuming the internal calibration is known. Under this assumption, 3D scene structures can be recovered up to a uniform scaling factor; this scaling parameter has an explicit physical meaning which is the distance of the first camera center from the 3D plane formed by the four points given in the correspondences. If this distance is known a priori, then a complete 3D Euclidean reconstruction can be obtained.

Faugeras and Lustman[FL88] showed that a homographic transformation between two cameras can be decomposed, and in general, there are two real solutions. In this paper, we propose a different way to decompose this matrix and perform a case by case analysis of different geometric situations. Finally, a general analytical closed-form solution is developed which covers all the possible cases. This enables us to reconstruct the 3D scene and, owing to this closed-form solution, the computation is very inexpensive.

*This work was supported in part by the National Science Foundation under grant number CDA-8922572, by the Advanced Research Projects Agency (via TACOM) under contract DAAE07-91-C-RO35, and by Rome Labs under contract F30602-94-C-0042.

†This work was performed when Zhongfei Zhang was at Computer Vision Lab., Dept. of Computer Science, Univ. of Massachusetts, Amherst, MA 01003.

2 Closed-form Solutions to the Decomposition of Homography Matrix

Consider a planar 3D object under the general viewing configuration illustrated in Fig. 1 with relative motion $\boldsymbol{R}$ and $\boldsymbol{t}$ between two cameras. Let $\boldsymbol{p_i}$ and $\boldsymbol{p_i}'$ be the images of any 3D point $\boldsymbol{P_i}$ on a 3D plane π. It can be shown[Zha96] that $\boldsymbol{p_i}$ and $\boldsymbol{p_i}'$ are related by

$$k_i \boldsymbol{p_i}' = \boldsymbol{A}\boldsymbol{p_i} \tag{1}$$

where k_i is a scaling factor. $\boldsymbol{A}$ is the homography mapping matrix, which is a 3×3 matrix defined as[Zha96]

$$\boldsymbol{A} = \boldsymbol{R}(\boldsymbol{I} - \frac{\boldsymbol{t}\boldsymbol{n}^T}{H}) \tag{2}$$

where $\boldsymbol{I}$ is a 3×3 identity matrix. The plane π is referred to as the *reference plane*. Thus, given at least 4 pairs of coplanar correspondences $\boldsymbol{p_i} \Longleftrightarrow \boldsymbol{p_i}', i = 1, ..., n, n \geq 4$, then based on Eq. 1, $\boldsymbol{A}$ can be solved up to a scale factor λ, i.e. a normalized 3×3 matrix $\boldsymbol{A_0}$ can be obtained such that

$$\boldsymbol{A} = \lambda \boldsymbol{A_0} \tag{3}$$

Define:

$$\boldsymbol{t_0} \stackrel{def}{=} -\frac{\boldsymbol{t}}{H} \tag{4}$$

$$k^2 \stackrel{def}{=} \boldsymbol{t_0} \cdot \boldsymbol{t_0}, \quad k > 0 \tag{5}$$

$$p \stackrel{def}{=} \boldsymbol{n} \cdot \boldsymbol{t_0} \tag{6}$$

Now we have:

$$\lambda^2 \boldsymbol{A_0}^T \boldsymbol{A_0} = \boldsymbol{I} + \boldsymbol{n}\boldsymbol{t_0}^T + \boldsymbol{t_0}\boldsymbol{n}^T + k^2 \boldsymbol{n}\boldsymbol{n}^T \tag{7}$$

We first assume that $\boldsymbol{t_0}$ and $\boldsymbol{n}$ are not aligned. We will relax this assumption later on to discuss those special cases when these two vectors are aligned. By definitions of eigenvalues and eigenvectors, it is easy to determine that the first eigenvector and eigenvalue of $\lambda^2 \boldsymbol{A_0}^T \boldsymbol{A_0}$ is:

$$\boldsymbol{v_1} = \boldsymbol{t_0} \times \boldsymbol{n}, \quad \eta_1 = 1 \tag{8}$$

Since $\lambda^2 \boldsymbol{A_0}^T \boldsymbol{A_0}$ is symmetric, as indicated in Eq. 7, the other two eigenvectors must be on the plane perpendicular to $\boldsymbol{v_1}$. It follows that we can assume that the other two eigenvectors are in the form:

$$\boldsymbol{v_{2,3}} = a\boldsymbol{t_0} + b\boldsymbol{n} \tag{9}$$

Thus,

$$\lambda^2 \boldsymbol{A_0}^T \boldsymbol{A_0}(a\boldsymbol{t_0} + b\boldsymbol{n}) = \eta_{2,3}(a\boldsymbol{t_0} + b\boldsymbol{n}) \tag{10}$$

where $\eta_{2,3}$ are the two corresponding eigenvalues.

Substituting Eq. 7 into Eq. 10 yields:
If $p \neq -1$,

$$\eta_{2,3} = 1 + p + \frac{2k(p+1)}{-k \pm \sqrt{k^2 + 4(p+1)}} \tag{11}$$

and

$$\boldsymbol{v_{2,3}} = \frac{-k \pm \sqrt{k^2 + 4(p+1)}}{2k(p+1)}\boldsymbol{t_0} + \boldsymbol{n} \tag{12}$$

If $p = -1$,

$$\eta_2 = k^2, \quad \eta_3 = 0 \tag{13}$$

and

$$\boldsymbol{v_2} = \frac{1}{k^2}\boldsymbol{t_0} + \boldsymbol{n}, \quad \boldsymbol{v_3} = \boldsymbol{t_0} \tag{14}$$

To simplify the mathematical expressions, let us introduce the following notation:

$$\gamma, \theta \stackrel{def}{=} \frac{-k \pm \sqrt{k^2 + 4(p+1)}}{2k(p+1)} \tag{15}$$

Based on the solutions of the obtained eigenvalues, we now examine the various cases for the three eigenvalues. By the definitions of k and p in Eq. 5 and Eq. 6, it is obvious that the following two constraints are always valid:

$$k \geq p, \quad ||k|| \geq ||p|| \tag{16}$$

We also know that $\lambda^2 \boldsymbol{A_0}^T \boldsymbol{A_0}$ is a real symmetric matrix, and thus it always has real eigenvalues. Therefore, a third constraint is:

$$k^2 + 4(p+1) \geq 0 \tag{17}$$

Based on the above three constraints, it can be shown that $\eta_2 \geq \eta_1 = 1$ where equality holds iff $p = -2, k = 2$ or $p = -1, k = 1$; and $\eta_3 \leq \eta_1 = 1$ where equality holds iff $p = k$.
Note that those conditions that make the above inequalities become equalities violate our previous assumption that $\boldsymbol{t_0}$ and $\boldsymbol{n}$ are not aligned. Therefore, in the general case that the two vectors are not aligned, we have a strict inequality: $\eta_2 > \eta_1 = 1 > \eta_3$

2.1 General Case: $\eta_2 > \eta_1 = 1 > \eta_3$

Assume the three singular values of $\boldsymbol{A_0}$ are: $\rho_2' > \rho_1' > \rho_3'$ and their corresponding orthonormal eigenvectors are $\pm\boldsymbol{u_2}, \pm\boldsymbol{u_1}, \pm\boldsymbol{u_3}$. Since $(\lambda\rho_1')^2 = \eta_1 = 1$, the normalization factor can be determined as:

$$\lambda = \frac{1}{\rho_1'} \tag{18}$$

Now we normalize the other two singular values:

$$\rho_{2,3} \stackrel{def}{=} \lambda \rho_{2,3}' \tag{19}$$

By recalling the relationship between the singular values and the eigenvalues, we have: $\eta_{2,3} = {\rho_{2,3}}^2$, which leads to:

$$k = \rho_2 - \rho_3, \quad p = \rho_2 \rho_3 - 1 \tag{20}$$

Now define:

$$\mu \stackrel{def}{=} ||\boldsymbol{v_2}||, \quad \delta \stackrel{def}{=} ||\boldsymbol{v_3}|| \tag{21}$$

Substituting Eq. 12 and Eq. 20

$$\pm \boldsymbol{u_2} = \frac{\gamma \boldsymbol{t_0} + \boldsymbol{n}}{\mu}, \quad \pm \boldsymbol{u_3} = \frac{\theta \boldsymbol{t_0} + \boldsymbol{n}}{\delta} \tag{22}$$

Solving the above four sets of two equations, we finally have four solutions:

$$\boldsymbol{t_0} = \pm \frac{\mu \boldsymbol{u_2} - \delta \boldsymbol{u_3}}{\gamma - \theta}, \quad \boldsymbol{n} = \pm \frac{\gamma \delta \boldsymbol{u_3} - \theta \mu \boldsymbol{u_2}}{\gamma - \theta} \tag{23}$$

and

$$\boldsymbol{t_0} = \pm \frac{\mu \boldsymbol{u_2} + \delta \boldsymbol{u_3}}{\gamma - \theta}, \quad \boldsymbol{n} = \pm \frac{\gamma \delta \boldsymbol{u_3} + \theta \mu \boldsymbol{u_2}}{\gamma - \theta} \tag{24}$$

Recalling that the solutions should be "real" in the sense that they should satisfy the constraint of $\boldsymbol{P} \cdot \boldsymbol{n} > 0$, two of the above four solutions are easily removed. Based on the remaining two solutions, the rotation $\boldsymbol{R}$ (based on Eq. 2) and the translation $\boldsymbol{t}$ (based on Eq. 4) can be obtained:

$$\boldsymbol{R} = \lambda \boldsymbol{A_0} (\boldsymbol{I} + \boldsymbol{t_0} \boldsymbol{n}^T)^{-1}, \quad \boldsymbol{t} = -H \boldsymbol{t_0} \tag{25}$$

Up to this point, we have found the two closed-form analytical solutions to the decomposition of the matrix $\boldsymbol{A_0}$ for this case.

2.2 Case $k = -p = 1$

This case leads to a violation of our assumptions. The constraints $p = -1$ and $k = 1$ really mean that the camera translates along $-\boldsymbol{n}$ a distance H; this puts the center of projection of the camera exactly on the reference plane. Regardless of its rotation, it will always see the reference plane as a line, which violates our assumption that the image correspondences should not be collinear.

2.3 Case $k = p$

The constraint $k = p$ implies that $\boldsymbol{t_0} = k\boldsymbol{n}$. Substituting this into Eq. 7 yields:

$$\lambda^2 {\boldsymbol{A_0}}^T \boldsymbol{A_0} = \boldsymbol{I} + 2k\boldsymbol{n}\boldsymbol{n}^T + k^2 \boldsymbol{n}\boldsymbol{n}^T \tag{26}$$

Obviously, any two arbitrary vectors perpendicular to $\boldsymbol{n}$ are eigenvectors of $\lambda^2 {\boldsymbol{A_0}}^T \boldsymbol{A_0}$, and they both have corresponding eigenvalues equal to 1. Moreover, since $\lambda^2 {\boldsymbol{A_0}}^T \boldsymbol{A_0}$ is symmetric, the two eigenvectors should also be perpendicular to each other. Specifically, let us define $\boldsymbol{v_1}$ and $\boldsymbol{v_3}$ as the two eigenvectors, and η_1 and η_3 their corresponding eigenvalues[1]. Thus, we have,

$$\eta_1 = \eta_3 = 1$$

$$\boldsymbol{v_1} \perp \boldsymbol{v_3}, \quad \boldsymbol{v_1} \perp \boldsymbol{n}, \quad \boldsymbol{v_3} \perp \boldsymbol{n}$$

Therefore, the third eigenvector must be $\boldsymbol{n}$, and its corresponding eigenvalue is $(k+1)^2$, i.e.

$$\eta_2 = (k+1)^2, \quad \boldsymbol{v_2} = \boldsymbol{n}$$

Similar to the general case derivation, now k and p can be obtained as:

$$k = p = \rho_2 - 1 \tag{27}$$

Therefore, $\boldsymbol{n}$ can be solved for as:

$$\boldsymbol{n} = \pm \boldsymbol{u_2} \tag{28}$$

Since $\boldsymbol{n}$ has to satisfy the constraint $\boldsymbol{P_i} \cdot \boldsymbol{n} > 0$, only one of the above two solutions is valid. Consequently, $\boldsymbol{t_0}$ is:

$$\boldsymbol{t_0} = k\boldsymbol{n} \tag{29}$$

and $\boldsymbol{t}$ and $\boldsymbol{R}$ can also be uniquely determined from Eq. 25. The geometric explanation of this case is that the camera center first translates distance kH along direction $\boldsymbol{n}$, and then rotates $\boldsymbol{R}$. This is the first solution for the case of $\eta_2 > \eta_1 = \eta_3 = 1$.

2.4 Case $k = -p > 2$

The constraint $k = -p$ implies that $\boldsymbol{t_0} = -k\boldsymbol{n}$. Thus, Eq. 7 becomes:

$$\lambda^2 {\boldsymbol{A_0}}^T \boldsymbol{A_0} = \boldsymbol{I} - 2k\boldsymbol{n}\boldsymbol{n}^T + k^2 \boldsymbol{n}\boldsymbol{n}^T \tag{30}$$

Similarly, the three eigenvalues are:

$$\eta_2 = (k-1)^2 > \eta_1 = \eta_3 = 1$$

and the three eigenvectors are: $\boldsymbol{v_2} = \boldsymbol{n}$, and $\boldsymbol{v_1}$ and $\boldsymbol{v_3}$ are two arbitrary vectors satisfying:

$$\boldsymbol{v_1} \perp \boldsymbol{n}, \quad \boldsymbol{v_3} \perp \boldsymbol{n}, \quad \boldsymbol{v_1} \perp \boldsymbol{v_3}$$

[1] We follow the convention in the general case that the largest eigenvalue and its corresponding eigenvector use subscript 2, the middle one is subscribed by 1, and the smallest one by 3.

Thus, k and p are solved for as:

$$k = \rho_2 + 1, \quad p = -\rho_2 - 1 \tag{31}$$

and finally,

$$\boldsymbol{n} = \pm \boldsymbol{u_2}, \quad \boldsymbol{t_0} = -k\boldsymbol{n} \tag{32}$$

subject to the constraint $\boldsymbol{P}_i \cdot \boldsymbol{n} > 0$.
The geometric explanation of the case for $k = -p > 2$ is that the camera first translates along $-\boldsymbol{n}$ a disdance kH, and then rotates $\boldsymbol{R}$. Since in this case $k > 2$, the new camera position must be on the other side of the reference plane. This unique solution, together with the unique solution derived for the case $k = p$, constitutes the two solutions for the case $\eta_2 > \eta_1 = \eta_3 = 1$.

2.5 Case $k = -p < 2$

This case is very similar to the case of $k = -p > 2$ except the three eigenvalues are:

$$\eta_2 = \eta_1 = 1 > \eta_3 = (k-1)^2$$

and the three eigenvectors are: $\boldsymbol{v_3} = \boldsymbol{n}$ and $\boldsymbol{v_2}$ and $\boldsymbol{v_1}$ are two arbitrary vectors satisfying

$$\boldsymbol{v_2} \perp \boldsymbol{n}, \quad \boldsymbol{v_1} \perp \boldsymbol{n}, \quad \boldsymbol{v_2} \perp \boldsymbol{v_1}$$

Here k and p can be solved for as

$$k = \pm \rho_3 + 1, \quad p = \mp \rho_3 - 1 \tag{33}$$

depending on $k > 1$ or $k < 1$. Although $\boldsymbol{n}$ has a unique solution $\boldsymbol{n} = \pm \boldsymbol{u_3}$ subject to the constraint $\boldsymbol{P}_i \cdot \boldsymbol{n} > 0$, $\boldsymbol{t_0}$ has two solutions because k has two solutions: $\boldsymbol{t_0} = -k\boldsymbol{n}$. Therefore, $\boldsymbol{t}$ and $\boldsymbol{R}$ also have two solutions. Again, the camera first translates along the $-\boldsymbol{n}$ direction a distance of kH, and then rotates $\boldsymbol{R}$. If $k > 1$, the camera is on the other side of the reference plane w.r.t. the initial camera position; if $k < 1$, the camera is on the same side of the reference plane.

2.6 Case $k = -p = 2$

This constraint implies $\boldsymbol{t_0} = -2\boldsymbol{n}$. Thus, Eq. 7 becomes singular:

$$\lambda^2 \boldsymbol{A_0}^T \boldsymbol{A_0} = \boldsymbol{I} \tag{34}$$

Therefore, any three arbitrary vectors perpendicular to each other are eigenvectors, and the corresponding eigenvalues are: $\eta_2 = \eta_1 = \eta_3 = 1$. Thus, $\boldsymbol{n}$ can be any of the three eigenvectors. Consequently, $\boldsymbol{t}$ and $\boldsymbol{R}$ have infinitely many solutions. The geometry of this case is that the camera center after motion is exactly at the mirrored position of the initial camera position w.r.t. the reference plane.

2.7 Summary of all the cases in the three eigenvalues

There are four possible orderings of the three eigenvalues. The general case is that all the three are distinct, i.e. $\eta_2 > \eta_1 > \eta_3$. The solution space of this case is a general ellipsoid which has three different axes, corresponding to the three different eigenvalues. The solution space for the case $\eta_2 > \eta_1 = \eta_3 = 1$ is a symmetric, cylindrical ellipsoid. In other words, the intersection of the ellipsoid at any point on the longer axis is a circle. This is why the other two eigenvectors can be arbitrary. Similarly, the case of $\eta_2 = \eta_1 = 1 > \eta_3$, is also a symmetric, cylindrical ellipsoid, in which the shorter axis corresponds to η_3. Finally, the solution space of case $\eta_2 = \eta_1 = \eta_3 = 1$ is a sphere, and the three eigenvectors can be arbitrary; hence, the number of solutions is infinite.

3 3D Reconstruction

Assume that $\boldsymbol{P}$ is an arbitrary 3D point which may or may not be on the reference plane. Using our conventional notation, i.e. $\boldsymbol{P} = (X, Y, Z)^T, \boldsymbol{p} = (x, y, 1)^T, \boldsymbol{p}' = (x', y', 1)^T$, and letting $\boldsymbol{R} = (\boldsymbol{R_1}, \boldsymbol{R_2}, \boldsymbol{R_3})$, where $\boldsymbol{R_1}, \boldsymbol{R_2}, \boldsymbol{R_3}$ are the column vectors of the rotation matrix, we can solve for $\boldsymbol{P}$[Zha96]:

$$X = \frac{t_{0Z}(\boldsymbol{R_1} \cdot \boldsymbol{p}') - t_{0X}(\boldsymbol{R_3} \cdot \boldsymbol{p}')}{x(\boldsymbol{R_3} \cdot \boldsymbol{p}') - (\boldsymbol{R_1} \cdot \boldsymbol{p}')} xH \tag{35}$$

$$Y = \frac{t_{0Z}(\boldsymbol{R_2} \cdot \boldsymbol{p}') - t_{0Y}(\boldsymbol{R_3} \cdot \boldsymbol{p}')}{y(\boldsymbol{R_3} \cdot \boldsymbol{p}') - (\boldsymbol{R_2} \cdot \boldsymbol{p}')} yH \tag{36}$$

$$Z = \frac{t_{0Z}(\boldsymbol{R_1} \cdot \boldsymbol{p}') - t_{0X}(\boldsymbol{R_3} \cdot \boldsymbol{p}')}{x(\boldsymbol{R_3} \cdot \boldsymbol{p}') - (\boldsymbol{R_1} \cdot \boldsymbol{p}')} H \tag{37}$$

Since $\boldsymbol{R}$ and $\boldsymbol{t_0}$ have two solutions in general, the final reconstruction also has two solutions. The two solutions are both subject to only one unknown scaling factor H, which is the distance of the first camera center from the reference plane. Therefore, we conclude that given at least four coplanar correspondences in two externally uncalibrated cameras, we can recover any 3D point in Euclidean space up to two solutions with only one uniform scale factor, which is the physical distance of the center of the first camera from the reference plane. If this distance is known, any 3D point can be *completely* recovered up to two solutions.
Now we investigate the problem of how to robustly implement this reconstruction technique in real world imaging situations, i.e. in the presence of noise and measurement error. Here we will only

consider the effect of error in localizing the position of the image points (i.e. localization error). To overcome the effect of localization error, more than four coplanar correspondences are needed so that a least mean square technique can be applied to obtain an optimal solution, assuming the localization error distribution is Gaussian or uniform. We adopt a fairly conventional technique for obtaining a robust solution in the presence of localization error. A least median square algorithm is employed in a global search for a subset of this n point set such that the least mean square solution of this subset gives rise to a minimum residual error[Zha96].

4 Experiments

The investigation of the robustness of the performance of this algorithm is conducted using simulation. We randomly choose n points on a 3D plane, and another 5 points not on the plane. The camera plane (i.e. the image) is digitized to 500×500 resolution. The camera aspect ratio is set to be 1, and the focal length is set to be 1000 pixels. The $n+5$ points are imaged at two arbitrary poses. Fig. 2 shows the performance of this algorithm under different levels of noise added to the two image planes w.r.t. the number of planar points, n. Based on this simulation result, we conclude that as long as the number of coplanar correspondences is greater than 15, the performance of the algorithm is quite good, with a reconstruction error of less than 5% in the presence of localization errors as large as 5 pixels. This provides an empirical threshold; in practice, we do not need more than 15 coplanar points to estimate the homography matrix in general, and the reconstruction is very stable even up to Gaussian noise with 5 pixels standard deviation.

Fig. 3 shows one frame of a box sequence taken by a CCD camera. The image resolution is 256×242. The location of 30 points (marked $+$ and $\times$ in Fig. 3) in a world coordinate system was measured to an accuracy of 1 mm. The box was rotated about its central vertical axis, while the camera was kept stationary. The 30 points were tracked over 8 frames. Table 1 shows the reconstruction result of the 30 points with input of frame 2 and 6. In this experiment, we use the left face (with points marked as $+$) of the box in Fig. 3 as the reference plane. The first ten points were used to search for a best estimate of the normalized homography matrix $\boldsymbol{A_0}$. Based on this $\boldsymbol{A_0}$, the relative geometry between the two cameras was obtained. This was then used to reconstruct all of the 30 points in 3D Euclidean space in the first camera coordinate system, which is what Table 1 lists. On a DecStation 5000, the total time for a complete combinatorial search for the best solution was less than 1 second. Table 1 lists the reconstructed 3D coordinates for each point, and the relative deviations of these scaling factors as compared with the ground truth. As can be seen from Table 1, most of the values are very close to this ground truth value (within 5% error) with an average absolute relative deviation as 3.79%.

5 Optimality

The homography $\boldsymbol{A}$ has 8 independent degrees of freedom (DOFs). On the one hand, each correspondence in Eq. 1 contributes 2 equations. Thus, 4 correspondences are the least to solve for $\boldsymbol{A}$, and hence, the algorithm is optimal in terms of number of correspondences. On the other hand, in Eq. 2, $\boldsymbol{A}$ contains $\boldsymbol{R}$ (3 DOFs), $\boldsymbol{t}$ (3 DOFs), $\boldsymbol{n}$ (2 DOFs), and H (1 DOF), for a total DOFs of $9 = 8 + 1$. Since this algorithm recovers relative geometry and 3D scenes up to one scale factor, it is optimal in terms of the internal calibration assumtion. In other words, we have to assume known calibration. It is impossible to add any unknown internal parameters in order to give such a reconstruction[Zha96].

6 Conclusion

A 3D reconstruction algorithm based on homography mapping between two views is presented. An analytical closed-form solution based on decomposing the homography matrix is proposed, together with a case-by-case discussion of different eigenvalue orderings. The algorithm is shown to be optimal and, in general, there are two solutions with one scale factor in Euclidean space. In future work, we intend to more completely investigate the robustness of this method based on more extensive simulations and tests on real data for which ground truth information is more readily available.

Acknowledgement

This work was inspired by an enlightening talk with John Oliensis. Thanks also go to Robert Collins, Ed Riseman and Richard Weiss for their useful conversations. Rakesh Kumar helped provide experimental images and data.

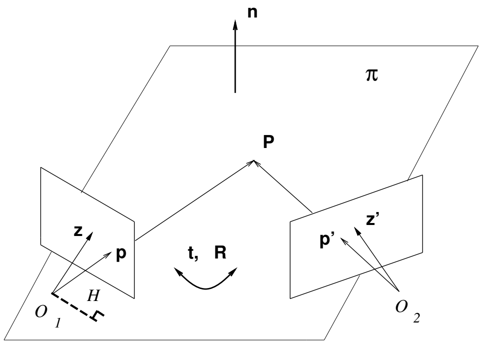

Figure 1: The geometry of a homography mapping.

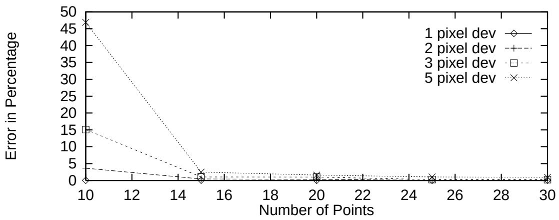

Figure 2: Error analysis results based on simulation

References

[BH87] S.D. Blostein and T.S. Huang. Quantization error in stereo triangulation. In *Proc. 1st ICCV*. IEEE, 1987.

[Fau92] O.D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proc. 2nd ECCV*, pages 563–578, Santa Margherita Ligure, Italy, May 1992. Springer-verlag.

[FL88] O.D. Faugeras and F. Lustman. Motion and structure from motion in a piecewise planar environment. *International Journal of Pattern Recognition and Artificial Intelligence*, 2:485–508, 1988.

[HGC92] R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *CVPR*. IEEE, 1992.

[Kum92] R. Kumar. *Model Dependent Inference of 3D Information From a Sequence of 2D Images*. Ph.D. Dissertation, COINS TR92-04, Univ. Mass., 1992.

[Sha93] A. Shashua. Projective depth: A geometric invariant for 3d reconstruction from two perspective/orthographic views and for visual recognition. In *ICCV*. IEEE, 1993.

[Zha96] Z. Zhang. *3D Reconstruction Under Varying Constraints on Camera Geometry for Robotic Navigation Scenarios*. Ph.D. Dissertation, CMPSCI, UMass, 1996.

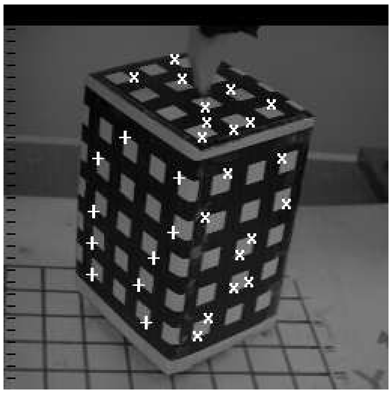

Figure 3: One frame of a box sequence.

Table 1

Reconstructed 3D Coordinates in Camera 1 System			Relative Deviations of the Recovered Scale Factors		
X	Y	Z	X	Y	Z
0.140	0.144	1.816	1.17%	5.03%	2.37%
0.028	0.063	1.707	-2.81%	3.33%	2.25%
0.202	0.112	1.912	3.20%	4.03%	2.77%
0.045	-0.050	1.771	-7.86%	-0.57%	2.67%
0.216	-0.001	1.979	2.97%	10.85%	2.96%
0.079	-0.105	1.844	0.20%	0.60%	2.76%
0.225	-0.073	2.019	3.04%	3.32%	3.32%
0.118	-0.155	1.910	-1.02%	4.24%	3.26%
0.233	-0.147	2.059	3.39%	3.78%	3.66%
0.097	-0.252	1.919	0.98%	2.86%	3.82%
0.055	0.334	1.981	-1.24%	-2.51%	2.50%
-0.006	0.125	1.688	-13.59%	0.94%	2.27%
0.124	0.267	1.886	6.41%	-0.06%	1.98%
-0.014	0.162	1.745	-13.25%	-1.36%	2.06%
-0.148	0.217	1.842	3.37%	-3.38%	2.42%
-0.073	0.153	1.733	3.44%	-1.44%	2.66%
-0.109	0.168	1.757	2.37%	-2.99%	2.73%
0.029	0.276	1.902	11.96%	-1.95%	2.37%
-0.019	0.197	1.795	-4.71%	-1.43%	2.27%
-0.056	0.244	1.865	8.04%	-1.78%	2.79%
-0.066	0.071	1.729	0.24%	1.37%	2.90%
-0.166	0.107	1.799	3.94%	1.60%	3.23%
-0.021	-0.016	1.757	4.13%	11.43%	2.51%
-0.187	0.004	1.884	3.59%	32.76%	3.84%
-0.077	-0.105	1.865	2.87%	3.07%	3.26%
-0.117	-0.057	1.859	1.84%	0.05%	3.75%
-0.069	-0.177	1.908	1.61%	4.83%	3.51%
-0.104	-0.172	1.923	1.25%	0.43%	4.08%
0.009	-0.273	1.897	8.69%	5.12%	3.87%
-0.032	-0.228	1.896	-8.46%	3.68%	4.06%