Linear Auto-Calibration for Ground Plane Motion

Joss Knight, Andrew Zisserman, and Ian Reid
Department of Engineering Science, University of Oxford
Parks Road, Oxford OX1 3PJ, UK
[joss,az,ian]@robots.ox.ac.uk

Abstract

Planar scenes would appear to be ideally suited for self-calibration because, by eliminating the problems of occlusion and parallax, high accuracy two-view relationships can be calculated without restricting motion to pure rotation. Unfortunately, the only monocular solutions so far devised involve costly non-linear minimisations which must be initialised with educated guesses for the calibration parameters. So far this problem has been circumvented by using stereo, or a known calibration object.

In this work we show that when there is some control over the motion of the camera, a fast linear solution is available without these restrictions. For a camera undergoing a motion about a plane-normal rotation axis (typified for instance by a motion in the plane of the scene), the complex eigenvectors of a plane-induced homography are coincident with the circular points of the motion. Three such homographies provide sufficient information to solve for the image of the absolute conic (IAC), and therefore the calibration parameters. The required situation arises most commonly when the camera is viewing the ground plane, and either moving along it, or rotating about some vertical axis. We demonstrate a number of useful applications, and show the algorithm to be simple, fast, and accurate.

1. Introduction

Camera calibration generally involves some compromise between automation and accuracy. Accurate calibrations can be obtained by imaging objects of known metric structure, such as in the method of Tsai [22], but greater accuracy usually requires greater user interaction. Since Tsai, there has been considerable research into reducing the level of metric information required without compromising accuracy. In particular, accurate planar objects are considerably easier to obtain than accurate regular solids, and Zhang [23] has shown how the known 3D structure required by Tsai can be reduced to known planar structure and unknown motion. Liebowitz and Zisserman [15] and Sturm and Maybank [19] use even weaker constraints on a planar object, such as known angles and length ratios.

Self-calibration relies only on the rigidity of the scene to extract the calibration, and must therefore process more image data to constrain the problem. However, it allows the process to become more easily automated. Indeed, it is the ability to post-process video sequences for the purposes of pose tracking or structure recovery that makes self-calibration essential in the special effects industry (*eg.* [1]).

In between the two extremes are a variety of situations in which some knowledge about the scene or motion can allow the calibration process to be reasonably accurate, while considerably more rapid and flexible than classical methods. A typical example of a known motion method is the non-translating camera algorithm first identified by Hartley [10]. This algorithm exploits the good approximation to pure rotation about the optical centre of many typical motions, such as that made by a camera on a pan-tilt unit.

The most common case of a non-general scene is a planar structure. Planar scenes have an advantage over pure rotations in general scenes because they provide equally strong matching constraints, yet allow a translational component to the motion; the distinct disadvantages of pure rotation are the large number of potential degeneracies and near-ambiguities [12], many of which can be resolved by including translation.

The disadvantages of general motion in an unknown planar scene are that each pair of images in a sequence only provides a small number of constraints on the calibration, and applying those constraints is non-trivial. Triggs [20] showed how to do this by using a version of the projection constraint of the absolute quadric [21]. His method is extremely flexible, but it is an iterative solution, with no closed-form counterpart available to initialise the optimisation. A number of linear solutions have been provided for the stereo case [5, 13, 14]. However, this work represents the first closed-form monocular calibration for unkown planar scenes that includes translation.

In this work we exploit some knowledge of the scene (that it is planar), and some knowledge of the motion. We show that when the camera moves about an axis normal to the

scene, a simple, linear calculation will provide the calibration from just three such motions, in what is essentially a closed-form solution for Triggs [20]. This commonly occurs when the scene being viewed is the ground plane. We are particularly interested in controlled calibration, especially for mobile robotics applications. Consequently we look at the two main scenarios: (i) a pan-tilt camera rotating about a vertical pan axis, and (ii) the camera (or scene) moving along the ground plane. In the latter case, the camera may be on a mobile robot, or simply on a tripod which is manually repositioned for each image.

1.1. Calibration using the absolute conic

This paper will use the common notational conventions as established in [11]. In particular, vectors appear in bold face ($\mathbf{x}$, $\mathbf{\Pi}$), and matrices in teletype (H, K). In addition, the camera internal calibration matrix K, with its components of focal lengths, principal point, and skew, is defined as

$$\mathtt{K} = \begin{bmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} .$$

The plane at infinity $\mathbf{\Pi}_\infty$ is the plane containing the points at infinity (directions) and lines at infinity which are the pre-images of vanishing points and vanishing lines respectively. It is invariant to translation (as the moon appears to be, for instance). Faugeras *et al.* [8] introduced the concept of the *absolute conic* into computer vision, which is a point conic on $\mathbf{\Pi}_\infty$ that is also invariant to rotation as a set. Its special property is that it projects to the *image of the absolute conic*, or IAC, $\boldsymbol{\omega} = \mathtt{K}^{-1}\mathtt{K}^{-\top}$, which depends only on the internal parameters. Finding this conic, or more usually its dual the DIAC, $\boldsymbol{\omega}^* = \mathtt{K}\mathtt{K}^\top$, is equivalent to calibrating for these parameters (since K can be recovered via Cholesky decomposition). The IAC and DIAC are point and line conics respectively, and are imaginary in that no real points lie on them.

Most calibration algorithms directly or indirectly find $\boldsymbol{\omega}$. One advantage of direct calculation of the IAC is that it is a simple matter to apply certain common constraints. As shown by de Agapito *et al.* [3, 4], the conditions on the elements of the IAC under the constraints of interest to us are:

1. *Zero skew*. If $s = 0$ then $\boldsymbol{\omega}_{12} = \boldsymbol{\omega}_{21} = 0$.
2. *Square pixels*. If $s = 0$ and $\alpha_u = \alpha_v$ then $\boldsymbol{\omega}_{11} = \boldsymbol{\omega}_{22}$.
3. *Fixed principal point*. If $s = 0$, $u_0 = 0$, and $v_0 = 0$, then $\boldsymbol{\omega}_{13} = 0$ and $\boldsymbol{\omega}_{23} = 0$.
4. *Focal length only unknown*. In this case, all the above constraints apply.

Just two constraints are required to calculate the focal length alone. Non-unity aspect ratios, and other principal point positions, can be taken care of by renormalisation of the image data or by image homographies.

2. Algorithm details

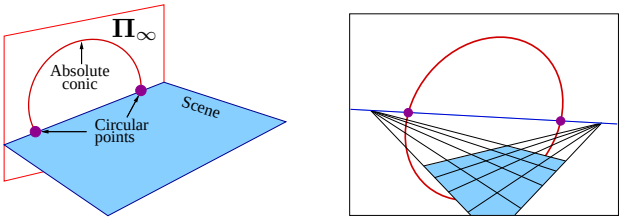


Figure 1. The circular points in a planar scene (left), and their images (right).

Both the methods of Triggs [21] and Demirdjian *et al.* [5] are effectively using the same basic algorithm, calculating the images of the scene plane's circular points. These are the (complex) points of intersection of the scene plane with the absolute conic on the plane at infinity $\mathbf{\Pi}_\infty$(see Figure 1). Since the images of these circular points lie on the image of the absolute conic, five or more such points are sufficient to calculate the IAC, and so the calibration. In both methods, only one conjugate pair of circular points need be explicitly calculated: the others are related to them by plane-induced homographies calculated from the images. The way the two methods differ is in the detection of this initial pair of points; Demirdjian, with the advantage of stereo imaging, uses a projective reconstruction to calculate all the scene horizons, which provide cubic constraints on the circular points; Triggs must make an iterative search.

The accuracy of both these methods is reliant on obtaining as many images as possible from widely separated viewpoints. This is not so much a deficiency in the algorithms as it is a reflection of the severe lack of constraints provided by a planar scene. In this work we show that a motion parallel to the scene allows for easy extraction of the circular points, and also improves accuracy from a minimal image set.

First we will derive the algorithm from a purely mathematical perspective. Following this we provide a complementary geometric argument.

2.1. Derivation

A camera with fixed internal parameters K, imaging a scene plane with normal $\mathbf{n}$ at distance d, undergoes rotation R and translation $\mathbf{t}$. The relationship between image points $\mathbf{x}$ before the motion and $\mathbf{x}'$ after is then a plane-induced homography H, *ie.* $\mathbf{x}' = \mathtt{H}\mathbf{x}$. In general a homography has three distinct fixed points corresponding to the three eigenvectors of the matrix. It will be shown that if the rotation axis is perpendicular to the scene plane, then there are two complex eigenvectors and these lie on the IAC regardless of the translation.

The homography $\mathtt{H}$ is given by

$$\begin{aligned}\mathtt{H} &= \mathtt{K}\mathtt{M}\mathtt{K}^{-1} \\ &= \mathtt{K}\left(\mathtt{R} + d^{-1}\mathbf{t}\mathbf{n}^{\top}\right)\mathtt{K}^{-1}, \end{aligned} \quad (1)$$

which is the well-known equation derived by Faugeras and Lustman [6]. The rotation matrix $\mathtt{R}$ has one real eigenvector $\mathbf{v}$, and two complex conjugate eigenvectors $\mathbf{e}$ and $\mathbf{e}^*$ with eigenvalues $\lambda_{\mathbf{e}}$ and $\lambda_{\mathbf{e}^*}$. $\mathbf{v}$ and the complex eigenvectors are orthogonal, since $\mathbf{v}$ represents the direction of the rotation axis, and $\mathbf{e}$ and $\mathbf{e}^*$ the plane perpendicular to it (which is invariant to $\mathtt{R}$). Consequently $\mathbf{v}^{\top}\mathbf{e} = \mathbf{v}^{\top}\mathbf{e}^* = 0$. The eigenvectors $\mathbf{e}$ and $\mathbf{e}^*$ are imaged as the points $(\mathtt{K}\mathbf{e})$ and $(\mathtt{K}\mathbf{e}^*)$ respectively, and lie on the IAC [11].

In the case that the rotation axis and scene plane are perpendicular, we have $\mathbf{n} = \mathbf{v}$. If we multiply $\mathtt{M}$ by $\mathbf{e}$, we have

$$\begin{aligned}\mathtt{M}\mathbf{e} &= \left(\mathtt{R} + d^{-1}\mathbf{t}\mathbf{n}^{\top}\right)\mathbf{e} \\ &= \mathtt{R}\mathbf{e} + d^{-1}\mathbf{t}\mathbf{v}^{\top}\mathbf{e} \\ &= \mathtt{R}\mathbf{e} + d^{-1}\mathbf{t}.0 \;=\; \mathtt{R}\mathbf{e} \;=\; \lambda_{\mathbf{e}}\mathbf{e}\ . \end{aligned} \quad (2)$$

This shows that $\mathbf{e}$ is an eigenvector of $\mathtt{M}$, and therefore $(\mathtt{K}\mathbf{e})$ is an eigenvector of $\mathtt{H}$. The same argument applies to $\mathbf{e}^*$. In this special case, the complex eigenvectors of $\mathtt{H}$ are the images of the eigenvectors of $\mathtt{R}$ and thus lie on the IAC regardless of the translational motion.

The complex eigenvectors are identical to those obtained by a pure rotation (i.e. with $\mathbf{t} = \mathbf{0}$), which corresponds to the *infinite homography*, $\mathtt{H} = \mathtt{K}\mathtt{R}\mathtt{K}^{-1}$, equivalent to a scene plane at infinity, $d = \infty$. Note, in the case of general motion where the rotation axis is not perpendicular to the scene plane, then for a non-zero translation the complex eigenvectors do not lie on the IAC in general.

2.2. Geometric argument

A plane intersects the plane at infinity in a line $\mathbf{L}$. The plane's circular points are the intersection of $\mathbf{L}$ with the absolute conic. As Figure 2 shows, all parallel planes intersect on $\boldsymbol{\Pi}_{\infty}$ at the same $\mathbf{L}$, and therefore have the same circular points. Consider the action of rotating the plane about its normal followed by a translation. Both the rotation and translation leave the plane orientation, and therefore the intersection line $\mathbf{L}$ unchanged, so that $\mathbf{L}$ is invariant (as a set). While in general points on $\mathbf{L}$ move along the line, the circular points are fixed because they are defined by the intersection with the absolute conic, which is itself invariant (as a set) to a Euclidean motion.

Consider now applying the action to a camera viewing the plane. Since the two points are fixed they have the same image before and after the motion and consequently are fixed points of the homography. Their image clearly lies on the IAC, the image of the absolute conic.

The locus of all points in 3-space for which $\mathbf{x} = \mathbf{x}'$ (i.e. have the same image before and after the motion) is known as the *horopter* curve. Generally this is a twisted cubic curve in 3-space passing through the two camera centres [17]. A twisted cubic intersects a plane in three points, and these are the three fixed points of the homography induced by that plane. In the case of plane-normal motion, two of the intersection points are at infinity where the twisted cubic, the scene, and $\boldsymbol{\Pi}_{\infty}$ coincide.

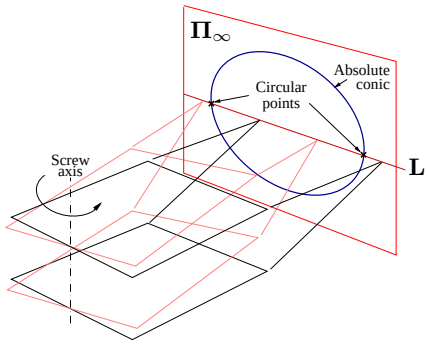


Figure 2. Parallel planes intersect on $\boldsymbol{\Pi}_{\infty}$ in a line $\mathbf{L}$. $\mathbf{L}$ is fixed if the planes rotate about a perpendicular axis and/or translate.

2.3. Algorithm

This method can be implemented in a number of ways. At one extreme, we can obtain a single pair of circular points from a single plane-normal motion, and transfer those points into other images using general plane-induced homographies. This, however, is quite heavily reliant on the quality of those original points. Further special motions will produce further point pairs to better constrain the problem.

At the other extreme, then, every motion made is a rotation about a plane-normal axis, so every homography calculated produces circular points that contribute to the calculation of the IAC. The minimal image set for a full calibration (all five parameters unknown) in this case becomes six images, instead of four. However, this algorithm does make the problem better conditioned, with every image contributing equally to the constraints, and errors.

The chosen algorithm can then be stated as follows for scenes containing horizontal planes (generally the ground plane):

To calibrate a camera from a horizontal plane

1. With the camera at a different 'attitude' for each motion, calculate plane-induced homographies $\mathtt{H}$ to motions about some vertical axis when viewing the scene plane.
2. Eigendecompose and find the circular points of each $\mathtt{H}$, their complex conjugate eigenvectors.
3. Fit the IAC to the points using the Direct Linear Transform [11]. Cholesky decompose its inverse (the DIAC) to give $\mathtt{K}$, the calibration matrix.

'Attitude' refers to the camera's orientation with respect to the scene plane at the start of each motion. For a pan-tilt unit or standard tripod this means varying elevation and/or cyclorotation. The orientation of the rotation axis with respect to the camera must change in order to provide different circular points for each motion.

The standard constraints mentioned in §1.1 (zero skew, square pixels *etc.*) can be applied during calculation of the IAC. Each motion provides two constraints, so in the general case the minimal set of motions is three. However, if zero skew is applied, only two motions are required, and if the principal point is also known, we only require a single pair of images.

2.4. Planar motion and plane-normal motion

Any rigid body motion is equivalent to a rotation about and translation along some axis in space, the screw axis. The special motion required by our algorithm requires only that the direction of the screw axis be normal to the scene plane. In most practically achievable circumstances, however, to achieve such a motion the translation, or pitch of the screw, will be zero. This is a *planar motion*, because the trajectory of the moving body lies in a plane.

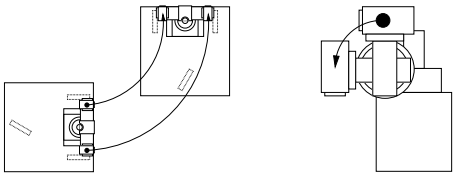


Figure 3. Planar motion is executed by, for instance, a robot moving along the ground plane (left), or an active head rotating about a single axis (right).

Planar motion occurs very commonly, as shown by Figure 3. We will be looking at cases in which the scene is the ground plane, and the rotation axis is vertical either because the camera is on a pan-tilt unit (which usually have vertical pan axes), or because the camera moves along the plane, or both.

Planar motion has been extensively studied by Armstrong and others [2, 7, 18], but the various planar motion calibration methods proposed are intended for general scenes and use the fundamental matrix. Our method bears more resemblance to other planar scene algorithms, and also in part to rotating camera algorithms, although they use the infinite homography, which provides additional constraints and cannot be applied here.

3. Experiments

3.1. Simulation

To assess the sensitivity of the algorithm to various parameters, it was tested over multiple runs on a simulated scene consisting of a cloud of random points on a plane in random position. The camera, which had a calibration varying around some nominal values, was placed at different orientations with the respect to the plane and rotated about an axis perpendicular to it. Except where the effect of additional motions was being tested, three motions were used in each case (the minimal set for to solve for all parameters), providing 3 homographies, and six circular points.

Three different implementations were tested in parallel. In the most basic, all calculations were kept entirely linear. In the second, the homographies were obtained using non-linear minimisation. In the third, the zero skew constraint was applied to the linear calibration. Finally for comparison, all homographies were assumed to have been derived from pure rotations, and a rotating camera algorithm [10] was used to provide the calibration. This was intended to show the conditions under which the use of a planar algorithm produces better results than assuming the camera is rotating. The results of the simulations are shown in Figure 4. Only focal length error is shown, since the other parameters tell a similar story.

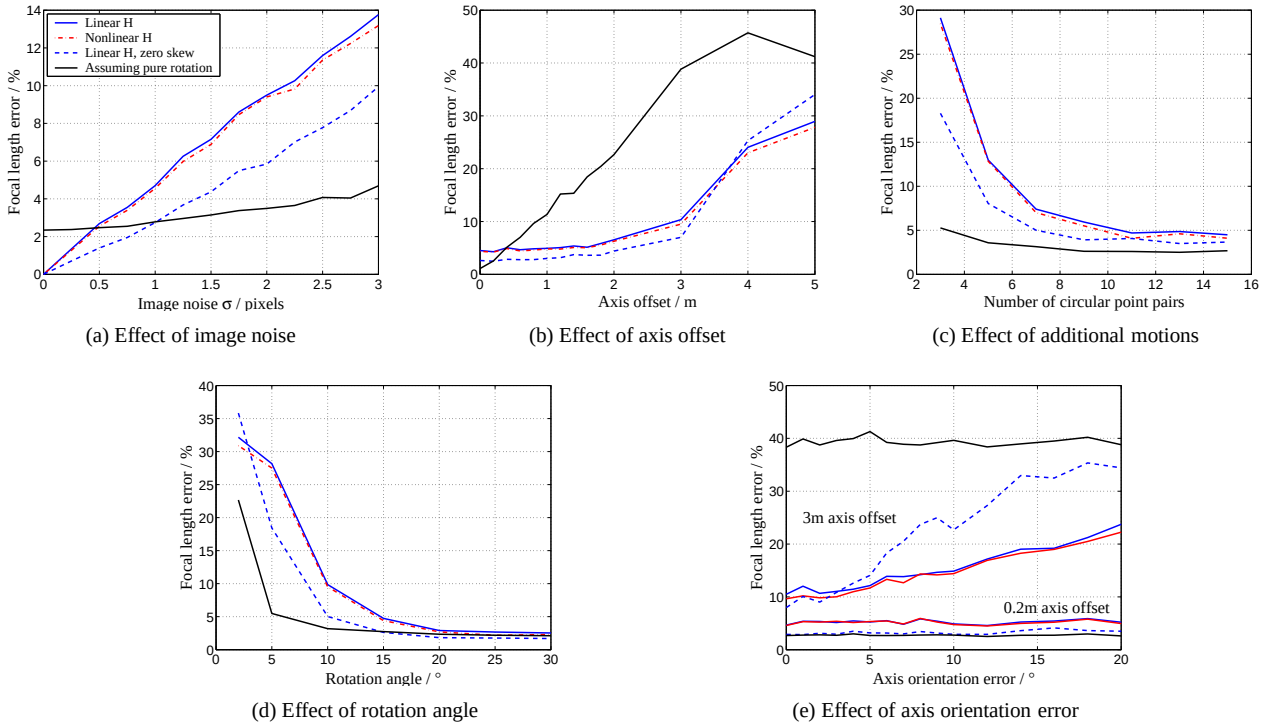
The simulations show that in a typical, properly controlled scenario, we should expect errors of around 5% in the focal length. It makes little difference whether or not the homographies are optimised, but application of the zero skew constraint provides great improvement.

The standard offset of the rotation axis from the camera was set to just 20cm, a typical value for the pan axis of a stereo pan-tilt unit. Since the scene itself was at a distance of 3m, the rotating camera method usually outperformed the planar calibration. There are, however, notable exceptions. For a start, the planar method reaches zero error when image noise is eliminated (Figure 4(a)), whereas the rotating camera method cannot due to the presence of translation in the motion.

Figure 4(b) shows that if the axis offset is increased, the planar method quickly becomes more accurate (in fact the transition occurs at around 40cm). This implies this method is best suited when the camera is translating along the ground plane and the position of the axis can be set at any distance from the camera (real tests will show that the calibration is in fact accurate in either case).

Figure 4(b) also shows that while the algorithm is insensitive to fairly large offsets, as the motion approaches pure translation (when the offset goes beyond about 3m, in this case), the errors increase dramatically. This effect disappears when there is zero noise, so clearly the calculated orientation of the axis becomes unstable in these circumstances. Such a situation would need to be avoided; however in a typical motion for which the scene is fixated, the rotation axis will lie within the scene, and will therefore not be distant from the camera.

Figure 4(c) tells us that additional motions can improve



(a) Effect of image noise (b) Effect of axis offset (c) Effect of additional motions

(d) Effect of rotation angle (e) Effect of axis orientation error

Figure 4. Results of the tests of the ground plane calibration algorithm on synthetic data. Nominal values for the parameters were 1 pixel noise σ, and 0.2m axis offset. The rotation angle was $15°$, except where a large number of motions or a large offset required it to be reduced to ensure the scene stayed in view. 200 matches were used for the calculation of each homography.

results, as would be expected, and Figure 4(d) indicates that larger rotations are preferable. Perhaps crucially, Figure 4(e) demonstrates that the algorithm can tolerate several degrees of error in the true orientation of the rotation axis: the pan axis of the camera platform, for instance, need not be precisely vertical for the results to be of use.

3.2. Real scene tests

There are a number of situations in which use might be made of this algorithm, and we address three of them here. In each case, the user has control over the motion itself, either physically or remotely, but not over the position or content of the scene plane. For these tests, the ground truth for the camera parameters was obtained to high accuracy using the method of Tsai [22].

First, the camera is on a passive tripod, and is calibrated from the ground plane by moving the tripod to different positions on the plane, altering the camera's orientation for each new motion (tripod calibration). In the second scenario, the camera is on an active head with a vertical pan axis (non-mobile head calibration). Finally, the camera is both active, and mounted on a mobile robot which moves along the ground plane (mobile head calibration). In each case the radial distortion was corrected prior to linear calibration. However tests showed that a single additional parameter in a minimisation for each plane-induced homography was sufficient to calculate and correct the distortion; alternatively it can be corrected with reasonable accuracy in closed-form by solving a quadratic eigenvalue problem, as demonstrated by Fitzgibbon [9].

3.2.1. Tripod calibration. Calibration from a tripod in any scene can be carried out using a rotating camera algorithm, but this does rely on the motions being a reasonable approximation to a pure rotation. Using the planar motion algorithm on the ground plane eliminates that concern, while still allowing for a fast result.

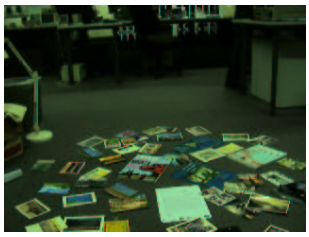
In our tests the ground plane was populated with some target objects, a camera orientation chosen, and then images captured before and after a motion of the tripod. Camera attitude was varied considerably, including cyclorotation, and some attempt was made to 'place' the rotation axis in a range of positions.

A total of 14 image pairs were taken, and the results are given in Figure 5(a). As can be seen, the results were extremely accurate, with only the principal point, which is anyway poorly constrained, not well estimated. In fact, a significant proportion of the 'minimal'[1] sets of images gave

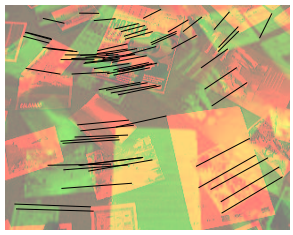
[1] 'minimal' is a misnomer, since the true minimal set for the least well-constrained method used (zero skew) is two image pairs.

Constraints	Percentage error		
	Focal length	Aspect ratio	P'pal pt
Focal length only	0.08	–	–
Fixed p'pal point	0.6	0.03	–
Square pixels	0.97	–	24.3
Zero skew	0.69	1.8	24.0

(a) Results from different constraints



(b) The scene



(c) Example image pair

Figure 5. (a) Tripod results when using all 14 circular point pairs. Principal point error is expressed as a percentage of focal length. (b) The scene and (c) one of the sets of point tracks, shown as combined images in separate colour channels (only about a third of the tracks are shown).

high accuracy results: the best set gives a focal length error of 0.1%. It appears that as long as a little effort is made to obtain pairs of images with a wide range of camera orientations and motions, accuracy can be reliably expected to be high even for small numbers of motions. The presence of objects not part of the dominant plane (such as in Figure 5(b)) were not a problem since they were ignored by the robust matching procedure.

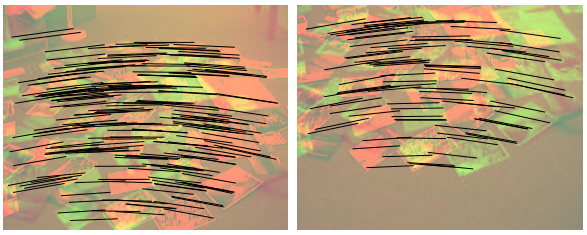
3.2.2. Non-mobile head calibration. The next test shows the viability of the use of this method for a straightforward pan-tilt unit with a vertical pan axis, such as a surveillance camera, or active head. In this case we made use of our stereo head, in which the pan axis is offset from the cameras by about 20cm. The calibration procedure was carried out entirely remotely.

In this case the elevation alone was varied to provide a new orientation for each pan motion. The vergence axis was fixed in the aligned position to simulate a standard monocular head. 13 motions about the vertical pan axis were made while viewing the same scene as for the tripod calibration, and the results are shown in Figure 6(a). In this case no combination of calculated circular points gave a focal length error much less than 10%, which is nevertheless good for a linear calibration. Possibly this is due to an orientation error in the pan axis, or the lack of variation in the cyclotorsional orientation of the camera.

For comparison, the image set was also tested using a rotating camera algorithm [10]. This method was expected to perform well because the offset of the pan axis from the camera is small, but it failed to get a focal length error under

Constraints	Percentage error		
	Focal length	Aspect ratio	P'pal pt
Focal length only	10.7	–	–
Fixed p'pal point	10.5	0.4	–
Square pixels	10.8	–	3.0
Zero skew	9.6	8.9	9.6

(a) Results from different constraints



(b) Two of the best 3 image pairs

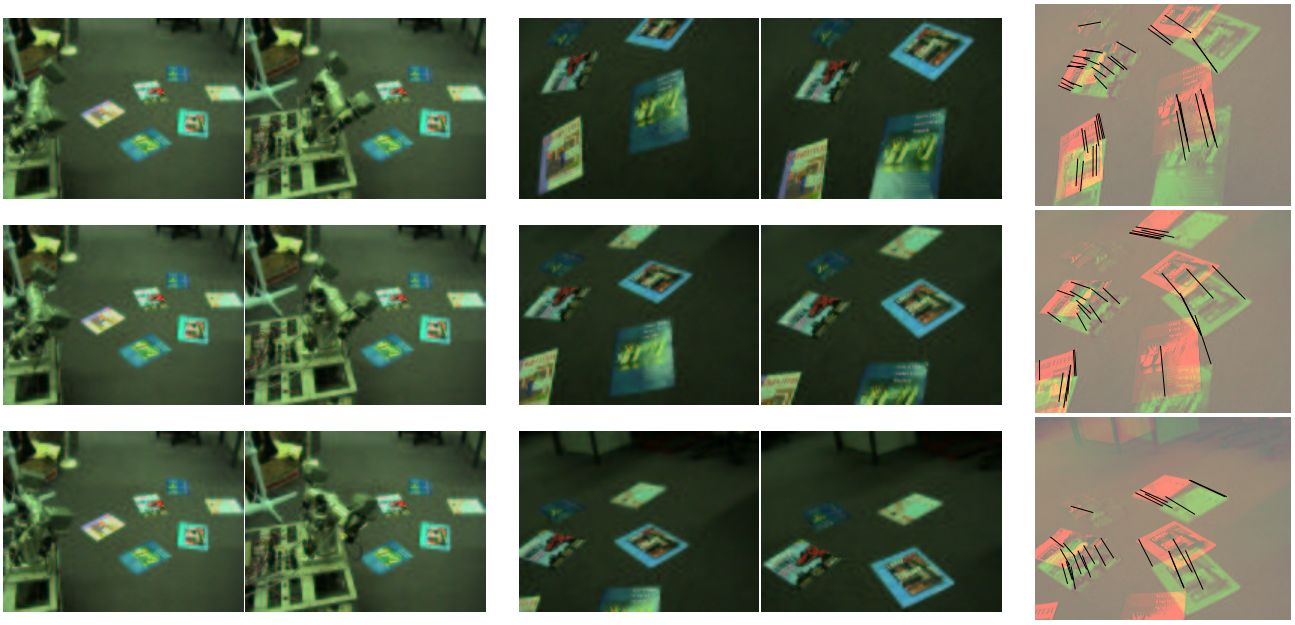
Figure 6. Non-mobile camera results, and two images from the best 'minimal' set in that sequence of 13 (the third was intermediate to the two shown).

62%. de Agapito *et al.* [3] make it clear why: this kind of pan-tilt motion is degenerate for the rotating camera model, with α_y and v_0 undetermined unless the square pixel constraint, at least, is applied. In this case we see that knowledge of the planarity of the scene has provided us with constraints that the rotating camera model cannot.

3.2.3. Mobile head calibration. Our stereo head is used as part of a visual navigation system, and is mounted on a mobile robot for that purpose. In the final test, the algorithm is used in a way that might be used in practice during navigation, to measure or correct the calibration of the cameras, which is essential for the stereo measurement process. As part of the experiment, the robustness of the algorithm was ascertained by making the minimal set of motions over a small range of orientations, and providing only sparse scene texture.

The situation is illustrated by the images of Figure 7. The robot paused in the process of normal navigation, and turned its head to obtain three images of the ground plane at different head orientations. It then moved forward a short distance and obtained three more images of the plane at the same vergence and elevation angles, but with increased pan to maintain overlap between images to be matched. The forward translation and the panning motion results in a rotation centre somewhere inside the scene.

Naturally this algorithm, like much of vision, requires some texture variation in the scene. However, plane-induced homographies can often be calculated to high accuracy even in texture-sparse scenes. This is illustrated by this test, with Figure 7(c) showing that considerably fewer matches were obtained in each image than for the previous experiments (between just 60 and 80 matches). However, as



(a) Robot positions (b) Image pairs (c) Point matches

Figure 7. Mobile active camera calibration: The robot (a) obtaining the images (b), which provided the correspondences (c).

Table 1. Calibrations from a mobile active camera platform.

Constraints	Percentage error		
	Focal length	Aspect ratio	P'pal pt
Focal length only	8.2	–	–
Fixed p'pal point	7.9	0.6	–
Square pixels	14.5	–	7.3
Zero skew	23.9	30.5	49.1

Table 1 shows, the resulting calibration was still quite accurate, with errors under 10% when the calibration is fully constrained, or the principal point is fixed. The results for zero skew were poorer, but it would always be advisable to apply constraints to standard values in such a restricted case.

As a final demonstration, the algorithm was applied to a suitable clip from a commercial film (Figure 8). In this crane shot, the camera pans, and translates both horizontally and vertically. From a single pair of images, solving for focal length only, we obtained a value of 2100 pixels. Using all the 3D structure (not just the ground plane) from the whole sequence of 340 images, some commercial software [1] gave a focal length of 2200 pixels, a difference of just 4.5%.

Using the plane-based linear calibration and the calculated homography, a rectified view of the ground plane was generated (see right-hand image of Figure 8), essentially a scene reconstruction. The annotations marking the paths show that their edges are now parallel, and something of the shape that the paths appear to form (a rectangle with a semicircular arch at the top). Considering the extreme perspective, this is a good achievement from just two input images.

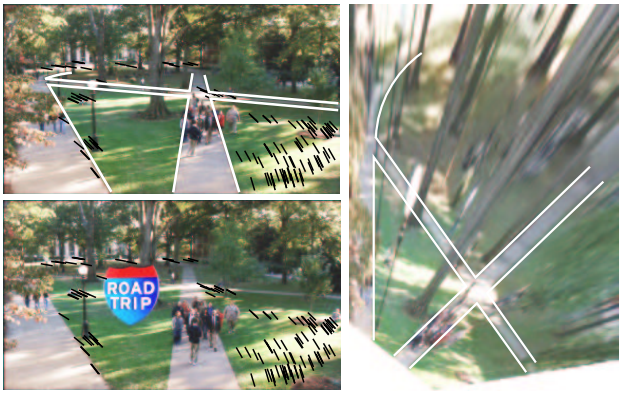


Figure 8. (Left) A pair of images from the film "Road Trip" for which a plane-normal motion, including vertical translation, was carried out. (Right) The rectified plane following calibration. The edges of the pathways are marked to emphasise the shape recovery. Full size colour images can be found at www.robots.ox.ac.uk/∼joss/cvpr2003.

3.2.4. Conditioning issues. It might be expected that in order for the calculation of the IAC to be well-conditioned, there should be an even spread of circular points around the centre of the image, which is the approximate centre of the IAC. This means a good variation in position and orientation of the scene horizon in the image, *ie.* a highly varied camera attitude, particularly cyclorotation.

Cyclorotation of the camera could only be achieved in the tripod tests, which did produce better results. However the other experiments show us that this conditioning issue can

be remedied by obtaining a sufficient quantity of high accuracy point matches. In the non-mobile head calibrations, for instance, the variation in elevation for the best minimal set was small, as illustrated by Figure 6(b).

4. Summary and Conclusions

In this paper we have described a linear algorithm for calibration from the ground plane, which can be used whenever motion along the ground plane, or rotation about a vertical axis, is possible. It is intended for use wherever camera motion can be controlled, but the scene itself cannot be easily manipulated or measured.

Tests have shown the algorithm to be fast, robust, and as long as the motions are chosen sensibly, accurate even in the minimal case. It has also been shown that a rotating camera algorithm, which would normally be the calibration method of choice for an active camera, is not always suitable. Our method can often fill in where the other fails. If it transpires that the camera is indeed rotating about its centre, this algorithm still works.

The ground plane, where it contains some texture variation, is an excellent source of accurate image transformations. Previous monocular planar calibration algorithms [16, 20] may well use the ground plane for this reason, but have no way of determining the necessary information to initialise their non-linear minimisations. Our algorithm provides a fast, linear result both usable by itself, and as a means of initialising a more computationally expensive optimisation. It also applies to a variety of other common real world situations, such as a fixed camera on a moving vehicle (a road car, for instance) when further constraints are used.

Acknowledgements

This work was supported in part by EU Project IST-1999-21125 (EVENTS).

References

[1] 2d3 Ltd., Boujou, 2000, http://www.2d3.com.

[2] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. In *Proc. 4th European Conf. on Computer Vision, Cambridge*, pages 3–16, 1996.

[3] L. de Agapito, R. I. Hartley, and E. Hayman. Linear self-calibration of a rotating and zooming camera. In *Proc. Computer Vision and Pattern Recognition Conference*, pages 15–21, 1999.

[4] L. de Agapito, E. Hayman, and I. Reid. Self-calibration of rotating and zooming cameras. *International Journal of Computer Vision*, 45(2), Nov. 2001.

[5] D. Demirdjian, A. Zisserman, and R. Horaud. Stereo autocalibration from one plane. In *Proc. 6th European Conference on Computer Vision, Dublin*, volume II, pages 625–639, 2000.

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

[7] O. Faugeras, L. Quan, and P. Sturm. Self-calibration of a 1D projective camera and its application to the self-calibration of a 2D projective camera. In *Proc. 5th European Conf. on Computer Vision, Freiburg*, volume I, pages 36–52, 1998.

[8] O. D. Faugeras, Q.-T. Luong, and S. J. Maybank. Camera self-calibration: Theory and experiments. In *Proc. 2nd European Conf. on Computer Vision, Santa Margharita Ligure, Italy*, pages 321–334, 1992.

[9] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proc. Computer Vision and Pattern Recognition Conference*, 2001.

[10] R. I. Hartley. Self-calibration of stationary cameras. *International Journal of Computer Vision*, 22(1):5–23, Feb. 1997.

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

[12] E. Hayman. *The Use of Zoom within Active Vision*. PhD thesis, Robotics Research Group, Oxford University Department of Engineering Science, Sept. 2000.

[13] J. Knight and I. Reid. Binocular self-alignment and calibration from planar scenes. In *Proc. 6th European Conference on Computer Vision, Dublin*, volume II, pages 462–476, 2000.

[14] J. Knight and I. Reid. Self-calibration of a stereo rig in a planar scene by data combination. In *Proc. 15th Int'l Conf. on Pattern Recognition, Barcelona*, volume 1, pages 411–414, 2000.

[15] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. In *Proc. Computer Vision and Pattern Recognition Conference*, pages 482–488, June 1998.

[16] E. Malis and R. Cipolla. Multi-view constraints between collineations: application to self-calibration from unknown planar structures. In *Proceedings of the 6th European Conference on Computer Vision, Dublin*, volume II, pages 610–624, 2000.

[17] S. J. Maybank. *Theory of reconstruction from image motion*. Springer-Verlag, Berlin, 1993.

[18] G. P. Stein, O. Mano, and A. Shashua. A robust method for computing vehicle ego-motion. In *IEEE Intelligent Vehicles Symposium (IV00), Dearborn, MI.*, Oct. 2000.

[19] P. Sturm and S. Maybank. On plane-based camera calibration: a general algorithm, singularities, applications. In *Proc. Computer Vision and Pattern Recognition Conference*, pages 432–437, June 1999.

[20] B. Triggs. Autocalibration from planar scenes. In *Proc. 5th European Conf. on Computer Vision, Freiburg*, volume I, pages 89–105, 1998.

[21] W. Triggs. Autocalibration and the absolute quadric. In *Proc. Computer Vision and Pattern Recognition Conference*, pages 609–614, 1997.

[22] R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, 1987.

[23] Z. Zhang. A flexible new technique for camera calibration. In *Proc. 7th Int'l Conf. on Computer Vision, Kerkyra, Greece*, pages 666–673, 1999.