Metric Rectification for Perspective Images of Planes

David Liebowitz and Andrew Zisserman
Robotics Research Group
Department of Engineering Science
University of Oxford
Oxford OX1 3PJ, UK.

Abstract

We describe the geometry, constraints and algorithmic implementation for metric rectification of planes. The rectification allows metric properties, such as angles and length ratios, to be measured on the world plane from a perspective image.

The novel contributions are: first, that in a stratified context the various forms of providing metric information, which include a known angle, two equal though unknown angles, and a known length ratio; can all be represented as circular constraints on the parameters of an affine transformation of the plane — this provides a simple and uniform framework for integrating constraints; second, direct rectification from right angles in the plane; third, it is shown that metric rectification enables calibration of the internal camera parameters; fourth, vanishing points are estimated using a Maximum Likelihood estimator; fifth, an algorithm for automatic rectification. Examples are given for a number of images, and applications demonstrated for texture map acquisition and metric measurements.

1 Introduction

It is well known that under perspective imaging a plane is mapped to the image by a plane projective transformation (a homography) [13]. This transformation is used in many areas of computer vision including planar object recognition [12], mosaicing [15], and photogrammetry [14]. The projective transformation is determined uniquely if the Euclidean world coordinates of four or more image points are known. Once the transformation is determined, Euclidean measurements, such as lengths and angles, can be made on the world plane directly from image measurements. Furthermore, the image can be rectified by a projective warping to one that would have been obtained from a fronto-parallel view of the plane (i.e. parallel to the image plane).

In this paper we show that it is not necessary to provide the Euclidean coordinates of four points; instead metric properties on the world plane, such as a length ratio and an angle, can be used directly to partially determine the projective transformation up to a particular (metric) ambiguity. This partial determination requires far less information about the world plane to be known, but is neverless sufficient to enable metric measurements of entities on the world plane to be made from their images.

Collins and Beveridge [2] made a significant step in this direction by showing that once the vanishing line of the plane is identified, the transformation from world to image plane can be reduced to an affinity. They used this result to reduce the dimension of the search, from eight to six, in registering satellite images. We improve on this result in four ways.

First, it is shown that by using known metric information the affinity can be reduced to a similarity. Section 2.2 describes three cases of providing this metric information from an image: a known angle, two equal though unknown angles, and a known length ratio. Each of these constraints can be represented simply as a circular constraint on two unknown parameters, and a closed form solution obtained by intersecting circles. This means that the problem of image registration considered in [2] can be reduced to a four dimensional search. Unlike the method of [2] no knowledge of the internal parameters of the camera is required. Faugeras *et al.* [7] used similar constraints in a 3D context, but only an iterative solution was given.

Second, in section 3, it is shown that an imaged plane can be rectified directly from metric information, without first identifying the vanishing line. This is illustrated for the metric constraints arising from right angles on the plane. Third, section 4, describes how the camera internal calibration parameters are constrained by the metric rectification of a plane. Fourth, vanishing points are estimated using a Maximum Likelihood Estimator (MLE) described in section 5.1. This substantially improves the accuracy of the results.

In man made scenes there are often two dominant orthogonal directions, for example aerial views of streets, interior

and exteriors of buildings. In this situation the vanishing points can be obtained automatically, and metric rectification is obtained up to a one dimensional ambiguity. This is described in section 5.3. The methods are demonstrated in section 6.

2 Stratification of projective rectification

As described by Koenderink [9] and Faugeras [6] metric structure recovery can be stratified so that first affine and then metric properties are determined. In this section we follow their approach.

Points on the image plane, $\mathbf{x}$, are related to points on the world plane, $\mathbf{x}'$, as $\mathbf{x}' = \mathtt{H}\mathbf{x}$, where $\mathbf{x}$ and $\mathbf{x}'$ are homogeneous 3-vectors. The projective transformation matrix $\mathtt{H}$ can be decomposed (uniquely) into a concatenation of three matrices, $\mathtt{S}$, $\mathtt{A}$ and $\mathtt{P}$, representing similarity, affine and 'pure projective' transformation respectively:

$$\mathtt{H} = \mathtt{SAP}$$

where

$$\mathtt{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{pmatrix}$$

and $\mathbf{l}_\infty = (l_1, l_2, l_3)^\top$ is the vanishing line of the plane. The vector $\mathbf{l}_\infty$ is homogeneous and has two degrees of freedom.

$$\mathtt{A} = \begin{pmatrix} \frac{1}{\beta} & -\frac{\alpha}{\beta} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Again, this matrix has two degrees of freedom represented by the parameters α and β. The geometric interpretation of these parameters is that they specify the image of the circular points [13]. The circular points $\mathbf{I}$ and $\mathbf{J}$ are a pair of complex conjugate points on the line at infinity which are transformed from co-ordinates $(1,\ \pm i,\ 0)^\top$ on the metric plane to $(\alpha \mp i\beta,\ 1,\ 0)^\top$ on the affine plane. Their significance lies in the fact that they are invariant to Euclidean transformations. Once they are identified metric properties of the plane are available. The final matrix in the decomposition is a similarity transformation

$$\mathtt{S} = \begin{pmatrix} s\mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{pmatrix}$$

where $\mathtt{R}$ is a rotation matrix, $\mathbf{t}$ a translation vector, and s an isotropic scaling. There are four degrees of freedom.

2.1 From projective to affine

The first stage is to determine $\mathtt{P}$, which requires identifying the vanishing line $\mathbf{l}_\infty$ of the plane. The vanishing line is the image of the line at infinity on the world plane. Parallel lines on the world plane intersect at vanishing points in the image, and the vanishing points lie on $\mathbf{l}_\infty$. Two or more such points determine $\mathbf{l}_\infty$. Once $\mathtt{P}$ is determined the image can be affine rectified (see figure 3), and affine properties such as length ratios on parallel line segments measured.

2.2 From affine to metric

It is assumed in the following that the geometry has been recovered up to an affine transformation by applying the matrix $\mathtt{P}$. Recovery of metric geometry requires an affine transformation of the plane, $\mathtt{A}$, that will restore angles and length ratios for non-parallel segments. In the following three methods of providing constraints on α and β are given. These are:

1. A known angle between lines;
2. Equality of two (unknown) angles; and,
3. A known length ratio.

In each case it shown that the constraint is a circle. This is in fact a circle in the complex plane since α and β are originally real and imaginary components, and the circles may be plotted on the plane with α as the real axis and β the imaginary. However, since α and β are real, the complex interpretation is not significant in seeking a solution.

Known angle:
Suppose θ is the angle *on the world plane* between the lines imaged as $\mathbf{l}_a$ and $\mathbf{l}_b$ (here lines $\mathbf{l}$ are homogeneous 3-vectors). Then it can be shown that α and β lie on the circle with centre

$$(c_\alpha, c_\beta) = (\frac{(a+b)}{2}, \frac{(a-b)}{2}\cot\theta)$$

and radius

$$r = |\frac{(a-b)}{2\sin\theta}|$$

where $a = -l_{a2}/l_{a1}$ and $b = -l_{b2}/l_{b1}$ are the line directions. Note, if $\theta = \pi/2$ the circle centre is on the α axis.

Equal (unknown) angles:
Suppose the angle *on the world plane* between two lines imaged with directions a_1, b_1 is the same as that between two lines imaged with directions a_2, b_2. Then it can be shown that α and β lie on the circle with centre on the α axis

$$(c_\alpha, c_\beta) = (\frac{a_1 b_2 - b_1 a_2}{a_1 - b_1 - a_2 + b_2}, 0)$$

and squared radius

$$\begin{aligned} r^2 &= (\frac{a_1 b_2 - b_1 a_2}{a_1 - b_1 - a_2 + b_2})^2 \\ &+ \frac{(a_1 - b_1)(a_1 b_1 - a_2 b_2)}{a_1 - b_1 - a_2 + b_2} - a_1 b_1 \end{aligned}$$

Known length ratio:
Suppose the length ratio of two non-parallel line segments

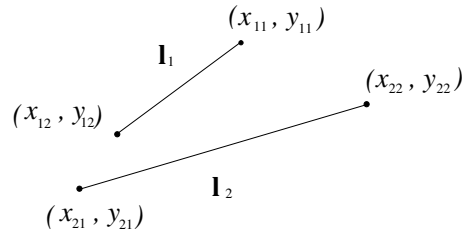


Figure 1. Notation for line segments in the known length ratio constraint.

is *s on the world plane*, and these line segments are imaged as shown in figure 1. Writing Δx_n for $x_{n1} - x_{n2}$ and similarly for y, it can be shown that α and β lie on the circle with centre on the α axis

$$(c_\alpha, c_\beta) = (\frac{(\Delta x_1 \Delta y_1 - s^2 \Delta x_2 \Delta y_2}{\Delta {y_1}^2 - s^2 \Delta {y_2}^2}, 0)$$

and radius

$$r = |\frac{s(\Delta x_2 \Delta y_1 - \Delta x_1 \Delta y_2)}{\Delta {y_1}^2 - s^2 \Delta {y_2}^2}|$$

2.3 Constraint combination

Two (independent) constraints are always required to determine α and β, and various combinations are illustrated in figures 4 and 5. It should be noted that the constraints are dependent on line orientation (in the affine rectified geometry), and the same constraint circle results from any parallel line sets.

Most of the examples employ right angles, since these are the most pervasive and useful angles in man made structures. For right angles the known angle constraint generates circles with centres on the α axis. Similarly, the known length ratio and equal (arbitrary) angle constraints also generate circles with centres on the α axis. Having all the constraint circles with centres on the same axis simplifies obtaining solutions because the constraint circles are symmetric with respect to this axis, and only intersections in the upper half plane need be considered.

3 Unstratified rectification

The previous section described a two step rectification process applying constraints sequentially on the projective and affine components of the rectification homography. It is possible, however, to determine the parameters of $\mathtt{AP}$ directly from metric information, without first using affine information, such as parallelism, to determine $\mathtt{P}$ from the vanishing line. In general direct application of the metric constraints generates non-linear constraints on the parameters. However, for orthogonal lines the constraint on the four rectification parameters is linear.

To obtain a linear constraint the parameters are represented by the conic $\mathtt{D}$ which is dual to the circular points. This conic is defined as $\mathtt{D} = \mathbf{I}\mathbf{J}^\top + \mathbf{J}\mathbf{I}^\top$ [16], and is represented by a rank two 3×3 matrix. Once the image of $\mathtt{D}$ is determined, the imaged circular points are also determined. From section 2, the circular points are imaged on the vanishing line at $((\alpha \mp i\beta)l_3,\ l_3,\ -\alpha l_1 - l_2 \mp i\beta l_1)^\top$ Consequently, once the circular points are determined the rectification parameters $\mathbf{l}_\infty$ and and α, β can be computed.

It can be shown that orthogonal lines are conjugate wrt $\mathtt{D}$, i.e. satisfy ${\mathbf{l}_a}^\top \mathtt{D} \mathbf{l}_b = 0$ for orthogonal lines $\mathbf{l}_a$ and $\mathbf{l}_b$. Each pair of orthogonal lines thus places a linear constraint on $\mathtt{D}$. Five orthogonal line pairs, i.e. five right angles, are sufficient to determine $\mathtt{D}$ linearly, provided lines of more than two orientations are included. Alternatively, $\mathtt{D}$ is determined by four orthogonal line pairs together with the rank two constraint, but the solution is non-linear.

4 Application to camera calibration

Caprile and Torre [1] have shown that three vanishing points for orthogonal directions allows partial calibration of a camera from a single view. The calibration method follows a construction showing that the principal point of the camera is at the orthocentre of the triangle which has the vanishing points as its vertices. This relies on prior measurement of the aspect ratio and the assumption that the image skew is zero.

Metric rectification of a plane also partially determines camera calibration. The five internal calibration parameters can be computed from the image of the absolute conic [5], and hence are determined once the imaged absolute conic is determined. The imaged circular points lie on the image of the absolute conic, and these points are known for a rectified plane. Each rectified plane then provides 2 points on the conic and thus 2 linear constraints on the conic. In the absence of additional constraints, 5 points, that is three non-parallel planes are required for camera calibration. If prior information about the camera is available, such as the aspect ratio or skew, only two non-parallel planes are required.

5 Implementation Details

5.1 Vanishing point estimation

A vanishing point is determined by the intersection of two imaged parallel lines. The intersection of the lines $\mathbf{l}_1$ and $\mathbf{l}_2$ is simply $\mathbf{x} = \mathbf{l}_1 \times \mathbf{l}_2$. Generally, there are more than two imaged parallel lines available and the vanishing point is thus over constrained.

The presence of measurement error ('noise') results in a set of line segments which do not intersect precisely in a point. A number of approaches to estimating the vanishing point have been proposed. A simple approach is the calculation of a weighted mean of all pairwise line intersections [1]. More elaborate has been the application of Bayesian statistics to error in projective spaces. Assuming a Bingham probability density function and mapping clusters of line intersection points to the unit sphere, a vanishing point is estimated as the point that minimises the sum of weighted orthogonal distances to the observed cluster [3].

In contrast to this, we define and implement a ML estimate of the vanishing point in order to minimize the errors where they occur: in the image. We have found empirically that this estimator significantly improves the accuracy of the metric rectification.

Suppose there are $n > 2$ line segments $\mathbf{l}_i$ and we seek to estimate the vanishing point $\mathbf{v}$. The ML estimate of the vanishing point $\hat{\mathbf{v}}$ involves finding also an estimate of the line segments $\hat{\mathbf{l}}_i$ such that $\hat{\mathbf{v}}$ lies on each line $\hat{\mathbf{l}}_i$ and the line set $\{\hat{\mathbf{l}}_i\}$ minimizes the Malhanobis distance from $\{\mathbf{l}_i\}$.

It is assumed that the error in the fitted line segments can be modelled by isotropic mean zero Gaussian noise on the end points. If the end points of $\mathbf{l}$ are $\mathbf{x}^a$ and $\mathbf{x}^b$, (figure 2), then the MLE minimizes

$$\mathcal{C} = \sum_i d_\perp^2(\hat{\mathbf{l}}_i, \mathbf{x}_i^a) + d_\perp^2(\hat{\mathbf{l}}_i, \mathbf{x}_i^b)$$

subject to the constraints $\hat{\mathbf{v}}.\hat{\mathbf{l}}_i = 0, \forall i$, where $d_\perp(\mathbf{x}, \mathbf{l})$ is the perpendicular image distance between the point $\mathbf{x}$ and line $\mathbf{l}$. An alternative cost function under different noise assumptions is given by Kanatani [8].

The cost function is minimized as follows: Given $\hat{\mathbf{v}}$ it can be shown that the cost $\mathcal{C}(\hat{\mathbf{v}})$ can be obtained in closed form. $\mathcal{C}(\hat{\mathbf{v}})$ can then be minimized over $\hat{\mathbf{v}}$ using the Levenberg-Marquart numerical algorithm [11]. An initial solution for $\hat{\mathbf{v}}$ is obtained from the null vector of the matrix $(\mathbf{l}_1, \mathbf{l}_2, \ldots \mathbf{l}_n)$ via singular value decomposition.

5.2 Image warping

Images are warped by applying the inverse homography to each pixel in the target image. The intensity at the source point in the original image is determining by bilinear interpolation. In order to automate the warping and ensure that the convex hull of the original image is correctly mapped into the rectangle of the target image, it is necessary to use oriented projective geometry [10].

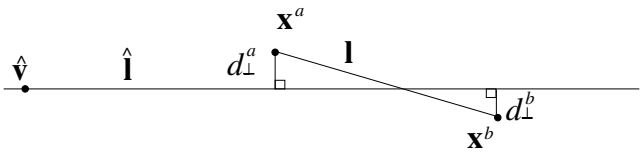


Figure 2. Geometry of the ML cost function.

5.3 Automatic detection of vanishing points and orthogonal directions

Automation of the correction process is achieved by detecting the two dominant directions of lines in the image, and assuming that these directions are orthogonal in the plane. The dominant directions are obtained from a frequency histogram on line direction (orientation), with the frequency weighted by segment length. Typically the histogram is bimodal, and is readily segmented. The lines in each dominant direction are assumed to be parallel and a vanishing point determined. The resulting affine image is likewise searched for dominant directions, and lines in the two directions constrained to be at right angles. Since this only provides one constraint on the affine parameters, there is an ambiguity in relative scale. Example histograms are given in figure 7. Note that while the histogram approach lends itself well to the detection of dominant directions of parallel lines in affine images, it is less well suited to vanishing point detection. The projective distortion of parallel lines results in distributed histogram maxima, such as a double peak in the presence of two widely spaced clusters of lines. More robust vanishing point detection may be accomplished with application of techniques such as the Hough Transform, as for example in [2] and [17].

Any point on the circle constraint chosen for the affine parameters will make those lines orthogonal, the choice of point determining the relative scaling in the two dominant directions. In the absence of any affine distortion, the parameters would have value $(0,\ 1)^\top$, so the point on the constraint circle closest to the point $(0,\ 1)^\top$ is chosen for the correction.

6 Examples and applications

Texture map acquisition. This example illustrates the variety of constraints available for metric rectification in the

case of a texture with rich geometry. Figure 3 shows the first stage where parallel lines are selected (manually) and used to determine the image of the line at infinity.

The three types of constraint and the resulting circles are illustrated in figure 4. Application of the direct (unstratified) method of section 3 using three of the light squares as constraints results in a metric image visually indistinguishable from figure 4(c). The angle between lines labelled 1a and 1b, for example, differs by 0.44^o between the metric images obtained using the stratified approach (shown) and a metric image obtained by the direct approach. Note that the orthogonality of these lines was used as a constraint in both cases. The rectified lines are not precisely orthogonal since both rectifications are overconstrained.

Using repeated elements. Repeated planar structure, where a feature appears at different orientations, allows use of the length ratio and equal angle constraints for unknown lengths and angles. Figure 5 (a) is an image of a rose window with twelve repetitions of the window segments. The line at infinity is found from parallel lines in the overall structure. The affine parameters are found, for demonstration purposes, only from a minimal set of constraints taking the indicated pairs of angles to be of equal magnitude. The resulting constraint circles and rectified image appear in figure 5 (b) and (c).

Euclidean measurements. Figure 6 demonstrates that Euclidean measurements may be made from an image once the homography is known up to a Euclidean ambiguity. The homography (PA) in this case is obtained from the orthogonality and length ratio of the inner goal area. Overall scaling (s), allowing true world measurement from the image, follows from setting the scaling for one of the known lengths in the image.

Automatic metric rectification. Figure 7 is rectified automatically using dominant line orientation in histograms. The histograms for both stages are shown, with clear separation of dominant direction peaks. Line orientation is computed over the range $-\pi/2$ to $\pi/2$ and 'wraps around', i.e. $-\pi/2$ and $\pi/2$ identify the same orientation.

7 Extensions

1. Lens distortion has been ignored throughout this paper. Where necessary the approach of Devernay and Faugeras [4] may be applied to remove radial distortion.
2. The symmetry of planar objects and circles can also be used to constrain rectification.
3. The MLE has so far been implemented for vanishing points from a set of line segments only. In cases where the vanishing line is over constrained, i.e. there are more than two vanishing points available, the line itself can be estimated using MLE.
4. Similarly, a ML estimator can be applied to the solution for (α, β) when more than two constraints are available.
5. The ideas developed in this paper can be extended to three dimensional structure, where metric rectification is of 3D projective structure obtained from uncalibrated images.

Acknowledgements

The authors would like to thank Dr Andrew Fitzgibbon for his invaluable assistance with software and many helpful discussions. We are grateful to EU Esprit Project Improofs and the University of the Witwatersrand Postgraduate Scholarship and Appeal Fund for financial support.

References

[1] B. Caprile and V. Torre. Using vanishing points for camera calibration. *IJCV*, pages 127–140, 1990.

[2] R. T. Collins and J. R. Beveridge. Matching perspective views of coplanar structures using projective unwarping and similarity matching. In *Proc. CVPR*, 1993.

[3] R.T. Collins. A Bayesian anlysis of projective incidence. In J.L. Mundy and A. Zisserman, editors, *Proc. 2nd European-US Workshop on Invariance, Azores*, pages 151–163, 1993.

[4] F. Devernay and O. Faugeras. Automatic calibration and removal of distortion from scenes of structured environments. In *SPIE*, volume 2567, San Diego, CA, July 1995.

[5] O. Faugeras, Q. Luong, and S. Maybank. Camera self-calibration: Theory and experiments. In *Proc. ECCV*, LNCS 588, pages 321–334. Springer-Verlag, 1992.

[6] O. D. Faugeras. Stratification of three-dimensional vision: projective, affine, and metric representation. *J. Opt. Soc. Am.*, A12:465–484, 1995.

[7] O.D. Faugeras, S. Laveau, L. Robert, G. Csurka, and C. Zeller. 3-D reconstruction of urban scenes from sequences of images. Tech. report, INRIA, 1995.

[8] K. Kanatani. Statistical optimization for geometric computation: theory and practice. Technical report, AI Lab, Dept of Computer Science, Gunma University, 1995.

[9] J. J. Koenderink and A. J. van Doorn. Affine structure from motion. *J. Opt. Soc. Am. A*, 8(2):377–385, 1991.

[10] S. Laveau. *Géométrie d'un système de N caméras. Théorie, estimation et applications.* PhD thesis, INRIA, 1996.

[11] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1988.

[12] C. Rothwell, A. Zisserman, J. Mundy, and D. Forsyth. Efficient model library access by projectively invariant indexing functions. In *Proc. CVPR*, pages 109–114, 1992.

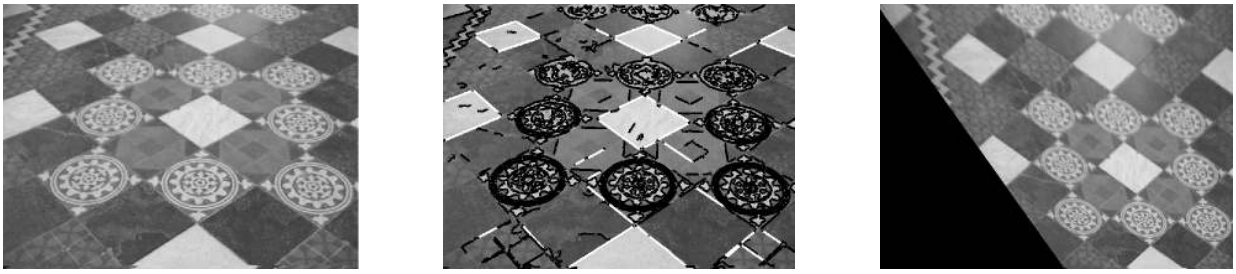
[13] J. Semple and G. Kneebone. *Algebraic Projective Geometry.* Oxford University Press, 1979.

[14] C. Slama. *Manual of Photogrammetry.* American Society of Photogrammetry, Falls Church, VA, USA, 4th edition, 1980.

[15] R. Szeliski. Image mosaicing for tele-reality applications. Technical report, Digital Equipment Corporation, Cambridge, USA, 1994.

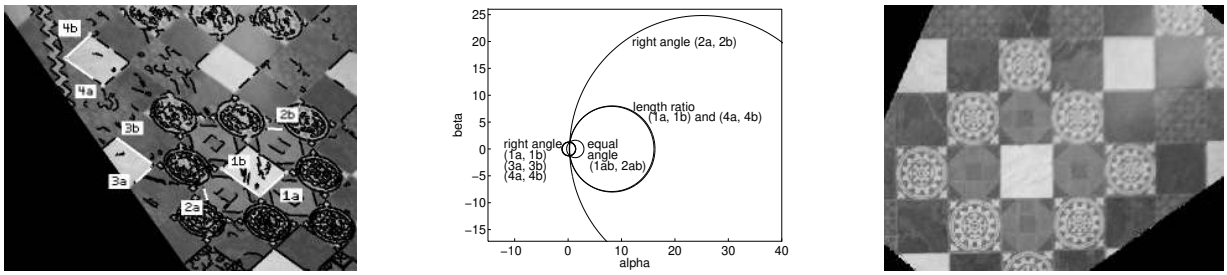
[16] W. Triggs. Autocalibration from planar scenes. In *Proc. ECCV*, 1998.

[17] T. Tuytelaars, L. Van Gool, M. Proesmans, and T. Moons. The cascaded Hough transform as an aid in aerial image interpretation. In *Proc. ICCV*, pages 67–72, January 1998.



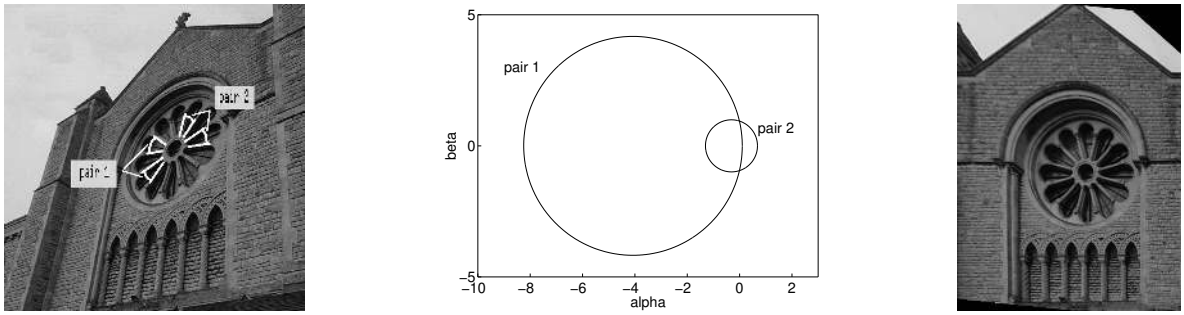
(a) (b) (c)

Figure 3. Affine rectification of a patterned floor. (a) The original image. (b) Line segments detected by Canny edge detection at sub-pixel accuracy; worm segmentation; and, fitting by orthogonal regression. Two parallel line sets (in white) intersect in two vanishing points, and determine the imaged line at infinity. The vanishing points are found by MLE. (c) The rectified affine image.



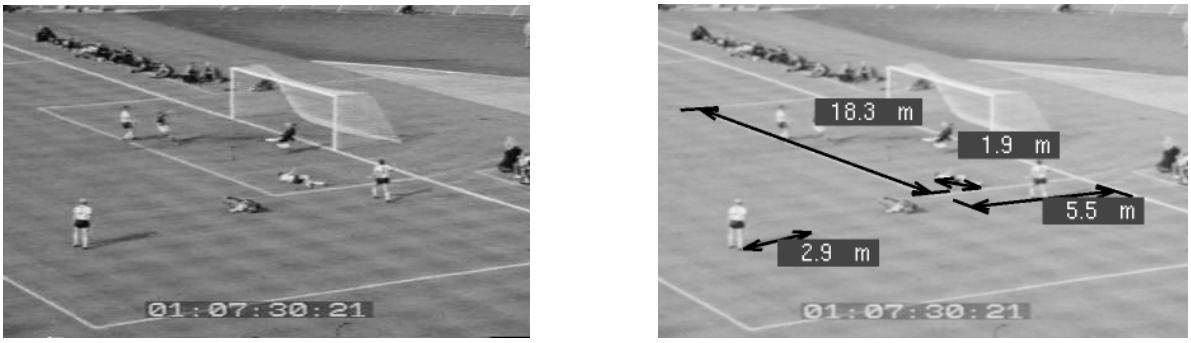
(a) (b) (c)

Figure 4. Metric rectification following figure 3(c). (a) Lines segments used to provide constraints. The lines shown in white are known to be in pairs at right angles. (b) Constraint circles in the $\alpha - \beta$ plane. The orthogonality of line pairs (1a, 1b), (3a, 3b) and (4a, 4b) gives the same circle (up to noise). The orthogonality of pair (2a, 2b) (which are not parallel to the first three) generates a different circle. The constraint of equal angle between pairs (1a, 1b) and (2a,2b) also generates a different circle. The constraint from unity length ratio of the sides of the squares bounded by (1a, 1b) and (4a, 4b) is also given. (c) The rectified metric image. Note that accuracy of the squares and circles demonstrate that a good estimation of the metric plane is achieved.



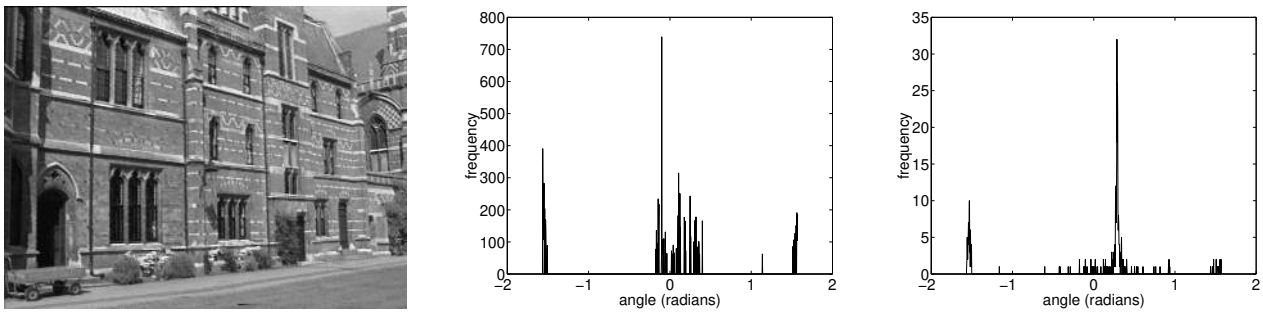
(a) (b) (c)

Figure 5. Rectification using repeated elements in a rose window. (a) The repetition of structure in the window allows application of the equal angles constraint for pairs 1 and 2. The true angle need not be known. (b) Constraint circles for the equal angle pairs. (c) The metric image.



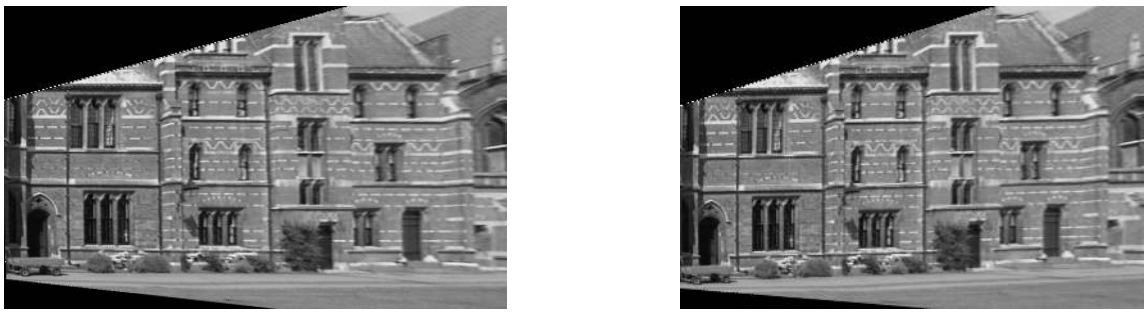
(a) (b)

Figure 6. Measurement of Euclidean quantities following perspective rectification. (a) Scene from an infamous England World Cup victory. Perspective correction using the parallel lines and known length ratio of the goal area, allows measurement in the plane from the perspective image. (b) Measurements of the goal area, a shadow and the approximate height of a fallen defender.



(a) (b) (c)

Figure 7. Bimodal histogram dominant direction detection. (a) Original image. (b) Orientation histogram of fitted lines in the original image. (c) Orientation histogram of fitted lines in the affine image (without edge length scaling). Note the sharpening around the modes.



(a) (b)

Figure 8. (a) Automated rectification. (b) Rectification with the correct relative scaling using a measured ratio of lengths of one of the windows. Note the slight difference in scaling.