An Automatic Approach for Camera Calibration from Vanishing Points

Lazaros Grammatikopoulos[1], George Karras[1], Elli Petsa[2]

[1] *Department of Surveying, National Technical University of Athens, GR-11364 Athens, Greece (lazaros@central.ntua.gr, gkarras@central.ntua.gr)*
[2] *Department of Surveying, Technological Educational Institute of Athens, GR-12210 Athens, Greece (petsa@teiath.gr)*

Corresponding Author (home address)
George Karras
Lerou 12, GR-11364 Athens, Greece,
Tel. +30 210 8622884, +30 210 7722685, fax. +30 210 7722677
e-mail: gkarras@central.ntua.gr

Abstract

Camera calibration is a fundamental task in photogrammetry and computer vision. This paper presents an approach for the automatic estimation of interior orientation from images with three vanishing points of orthogonal directions. Extraction of image line segments and their clustering into groups corresponding to three dominant vanishing points are performed without any human interaction. Camera parameters (camera constant, location of principal point, two coefficients of radial lens distortion) and the vanishing points are estimated in a one-step adjustment of all participating image points. The approach may function in a single-image mode, but is also capable of handling input from independent images (i.e. images not necessarily of the same object) with three and/or two vanishing points in a common solution. The reported experimental tests indicate that, within certain limits, results from single images compare satisfactorily with those from multi-image bundle adjustment.

Keywords: Automation, Calibration, Distortion, Edge, Extraction

1.   Introduction

Extraction and exploitation of the information, about both camera and object geometry, which is inscribed into the vanishing points of a perspective image is a significant topic of research, particularly within the context of automation. The use of vanishing points refers essentially to three fundamental – though not necessarily independent – tasks, those of camera calibration, image orientation and object reconstruction. In most cases uncalibrated images of man-made structures are involved, which usually lack conventional control information. Camera calibration is the particular topic of this contribution.

In certain instances the calibration step may be by-passed since, under circumstances, even one suitable vanishing point on an uncalibrated image might suffice for affine planar reconstruction and thus allow 1D measurement (Grammatikopoulos et al., 2002). For camera calibration, however, it is known that the primary elements of interior orientation (camera

constant and principal point location) may be estimated – along with the camera rotation matrix – from vanishing points of three orthogonal directions on a single image. Merritt (1958) and Gracie (1968) have provided equations for estimating interior orientation parameters and camera attitude from configurations with such vanishing points (image principal point is the orthocentre of the triangle formed by the vanishing points). Methods explicitly or implicitly using, in different conceptual frameworks, three image vanishing points of orthogonal directions for the purposes of camera calibration have been subsequently reported in both photogrammetry and computer vision (Bräuer-Burchardt and Voss, 2001; Caprile and Torre, 1990; Cipolla et al., 1999; Liebowitz et al., 1999; Sturm and Maybank, 1999; van den Heuvel, 1999).

Yet, the demand that an image displays three vanishing points which are suitable for reliable camera calibration (for instance, none is close to infinity) is admittedly not trivial. In contrast, images (e.g. of 2D objects) with pairs of finite vanishing points are much more common. Two vanishing points of orthogonal space directions may permit, assuming fixed principal point, to recover image rotations and camera constant (Karras et al., 1993) or directly perform planar rectification. If the principal point is not known, metric rectification requires additional information, e.g. a length ratio in different directions (Liebowitz et al., 1999), yet the principal point remains unknown.

However, although one single image with two vanishing points lacks the information for full camera calibration, the authors have shown that combinations of several independent (single) images with vanishing points in two orthogonal directions supply adequate information for estimating camera geometry, including aspect ratio (Grammatikopoulos et al., 2004). By 'independent' or 'single' images it is meant that the contents of images need not be the same as in other cases of calibration, i.e. no stereo views or block of images of the same object are required for camera calibration based on vanishing points. Using the relation of these two vanishing points to the camera parameters, the approach in the above publication rests on a direct geometric interpretation regarding the locus of the projection centre in the image system. These loci are hemispheres, with centres belonging to the image plane and diameters fixed by the two orthogonal vanishing points. This non-linear equation involves explicitly the camera interior orientation parameters along with the coordinates of the vanishing points.

Against this background, this paper presents an approach for the automatic calibration of single images with three vanishing points of orthogonal directions (which, at the same time, can perform a common calibration of several independent images with three and/or two vanishing points). Exploiting a priori information about object geometry – line parallelism and perpendicularity constraints – and assuming square pixels, the three primary parameters of interior orientation, namely camera constant and principal point location, are estimated along with two coefficients of radial symmetric lens distortion without any user interaction. Camera calibration is achieved iteratively through a simultaneous least-squares adjustment of all image points belonging to lines which intersect in the three dominant vanishing points of the scene, on the assumption that these correspond to orthogonal space directions. In this sense, the approach presented here relates particularly to those of Becker and Bove (1995), Bräuer-Burchardt and Voss (2001), and van den Heuvel (1999).

2. One-step camera calibration approach

Parameter computation in van den Heuvel (1999) is split into two steps. First, the parallelism constraint established among the normal vectors of the interpretation planes of parallel 3D lines allows estimation of vanishing points and coefficient k_1 of radial distortion, with the

image centre chosen as the centre of symmetry. In a subsequent step, the perpendicularity constraint is exploited, using observations adjusted from the previous step, for the estimation of camera parameters. The nominal camera parameters should be known, else user interaction is necessary. Bräuer-Burchardt and Voss (2001) first estimate the symmetric radial distortion (coefficients k_1, k_2 and location of the point of best symmetry of distortion) along with the vanishing points via an iterative process. Interior orientation parameters are then computed from the corrected vanishing points. In both approaches, detection of vanishing points and classification of line segments in orthogonal directions run automatically. On the other hand, Becker and Bove (1995) estimate radial and decentering image distortions by minimizing the dispersion of vanishing points on the Gaussian Sphere. The internal camera parameters are then recovered in a second independent step.

Here, unlike the above approaches, vanishing points and interior orientation parameters (including radial distortion) are estimated simultaneously. Of course, detection of vanishing points along with the estimation of radial lens distortion is in fact an entirely autonomous task; it yields vanishing points estimated as the common points of concurring lines free from distortion. The latter refers to the point of best symmetry of distortion S which, strictly speaking, does not coincide with the image principal point $P(x_o, y_o)$. The six image coordinates of the vanishing points of three orthogonal directions supply an equal number of equations which then yield, with no redundancy, the remaining group of unknown quantities, namely the three camera geometry parameters and the image rotation matrix. Generally, however, if not the highest possible accuracy is needed the principal point itself is mostly assumed as the zero-point of lens distortion ($S \equiv P$). In this case, the adjustment should be performed in a unified process, with x_o, y_o as common unknowns in the two groups of equations. In fact, given the accuracy expected when calibrating with vanishing points, it is questionable whether the point of best symmetry has here something substantial to offer – other than a distinction from the principal point to simply allow splitting the process in two steps.

Furthermore, a simultaneous adjustment is preferable if more than the minimum of three vanishing points are available. This would be the case, for instance, if additional directions in space are known. But a redundancy in vanishing points also exists when two or more independent images (from the same camera) with three vanishing points are at hand, since in this case all triplets of vanishing points should be constrained to define triangles reflecting the same camera geometry. In such a case, a unified approach is evidently advantageous against two successive adjustment processes. This is also the case when independent images (of the same or altogether different scenes) with only two vanishing points are adjusted, as outlined in Grammatikopoulos et al. (2004). Clearly, combinations of images with three and two vanishing points are also possible. Finally, the unified approach presented here is also useful when images of the same scene have been taken from closely neighbouring standpoints, and hence no bundle adjustment is feasible.

In principle, the location of the point of best symmetry S might also be included as unknown in the one-step algorithm. This is not the case here (radial distortion refers to the principal point x_o, y_o), while decentering lens distortion has also been neglected. Automation of the procedure means that a vanishing point detection phase is necessary before the adjustment, preceded by an edge extraction and linking process.

2.1.   Vanishing Point Detection

Different approaches have been reported as regards automatic detection of vanishing points. Barnard (1983) first introduced the use of the Gaussian Sphere as an accumulation space: every image line segment defines a corresponding major circle on the 'unit sphere', which is centred onto the optical centre; the intersection of two or more such circles defines their vanishing point. The search of such intersections is performed using Hough transform in the tessellated sphere and local maxima are detected as vanishing points. Main advantage of this bounded accumulation space is the representation of finite as well as infinite vanishing points. Yet, it has been pointed out that spherical sampling is here different in the x and y directions, generating non-uniform cells. Quan and Mohr (1989) extended this method using an accumulator generated by hierarchical division of the sphere, while Lutton et al. (1994) by-passed the problem via a combination of regular and irregular quantization. In the last case, errors due to the finite extension of the image, its quantization and the ambiguity of detected segments have been compensated. Shufelt (1999) introduced the 'interpretation plane swath model' to handle the orientation error in image line segments and the interpretation planes they produce. Magee and Aggarwal (1984) compute all intersections among pairs of line segments on the Gaussian Sphere, without using the sphere as an accumulator space. On this basis, Gallagher (2002) proposed an algorithm using conditional probabilities established by a ground truth training set. A comparison of parameterization models of the Gaussian Sphere with respect to noise is found in Wenzel and Grigat (2006).

Brillault and Mahony (1991) have presented a different accumulator space involving the uncertainty of lines to ensure a constant detection quality of vanishing points. The exploitation of the polar parameter space as accumulation space is also found in several publications (Straforini et al., 1993; Cantoni et al., 2001), while approaches for estimating vanishing points based on the 'expectation maximization' algorithm have also been reported (Antone and Teller, 2000; Kosecká and Zhang, 2002). Alternatively, vanishing points can be detected through texture analysis (Rasmussen, 2004). In the latter approach the perspective pose of textured planes is estimated by projecting spectral information onto the unit sphere.

The procedure adopted here consists of two basic steps, that of edge extraction and linking, and that of vanishing point detection.

2.1.1.   Edge extraction and linking

First, image edges are extracted with sub-pixel accuracy and connected to segments by means of polygonal approximation based on a suitable 'straight-line threshold' (Thormaehlen et al., 2003), which limits the maximal accepted distance of edge points from the line segment. The line parameters are estimated through least-squares line-fitting. A threshold on segment length allows excluding very short segments from the process. As outlined in section 2.2.3, the threshold size decreases after the first estimation of the radial lens distortion coefficients.

2.1.2.   Vanishing point detection

In this phase, line segments extracted from the previous step are classified into groups corresponding to dominant space directions. The method for vanishing point detection used here relies on van den Heuvel (1998) and Rother (2000), who employ the additional criterion of orthogonality of space line directions for establishing vanishing points of three mutually orthogonal directions. As suggested by Rother (2000), the image plane itself is best chosen as the accumulator space thanks to its property of preserving the original distances among points

and lines (translational and rotational invariance), while the intersections of all pairs of image line segments are selected as accumulator cells.

For each possible vanishing point (accumulator cell) Rother (2000) computes the contribution of every image segment s by means of a voting scheme (also adopted in Bauer et al., 2002). Votes are the sum of two weighted contributions. The first stems from the angle α between the line segment s and the line joining its midpoint M with the specific vanishing point V. The weighted contribution of all segments for which α does not exceed a threshold α_0 is inversely dependent on their value of α. The second weighted contribution is directly proportional to the length of the segment. Eventually, votes from all accepted segments for some vanishing point are added to yield the final 'strength' for this point. Empirical values for these two weighting factors are given in Rother (2002).

Here, the need for weighting factors (one for angles, one for segment lengths) in the voting process is eliminated to allow a simpler uniform treatment of long and short segments. This succeeds by introducing the uncertainty σ (standard error) of the slope angle of each line as estimated from the line-fitting adjustment. This uncertainty (illustrated in Fig. 1 together with segment s, its midpoint M, vanishing point V and angle α) may also be regarded as the uncertainty of α.
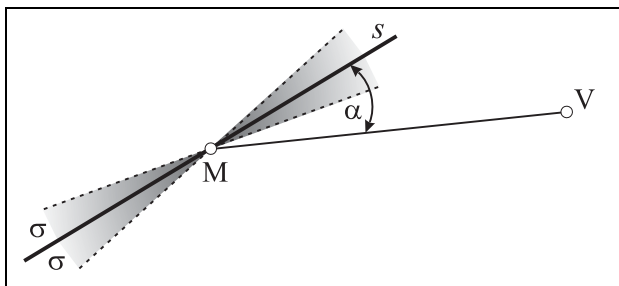

Figure 1. Angle α (formed by segment s and the line joining its midpoint M with potential vanishing point V) and its uncertainty σ.

Hence, a realistic estimation of angle α formed by segment s and line VM is $\alpha + \sigma$, namely the computed angle plus its uncertainty. The standard error σ, extracted from the covariance matrix of the line-fitting adjustment, reflects both the effect of random errors (deviations of points from the best-fitting line) as well as the effect of geometry of the adjustment, in which the length of the line segment plays the central role. Hence, the impact of segment length is incorporated in σ, combined at the same time with the standard error of image coordinates (since shorter segments do not necessarily produce larger uncertainties in slope and vice versa, depending on the quality of measured data). In this sense, there is no need for empirical weighting factors, and the vote v contributed by each segment for a particular vanishing point may be computed as

$$v = 1 - \frac{\alpha + \sigma}{\alpha_0} \tag{1}$$

with α_0 being a suitable threshold value. The votes from all segments with $(\alpha+\sigma) < \alpha_0$ for a particular vanishing point are accumulated, resulting in a total score for this potential

vanishing point. It might be observed here that, strictly speaking, the uncertainty of α is due not only to the orientation error of line segments but also to the uncertainty of their midpoint M; yet, this effect is regarded as negligible compared to the error of line slope.

Next, the accumulator cells along with their corresponding final scores are checked (following Rother, 2000) against certain geometrical criteria for the determination of the three dominant vanishing points of mutually orthogonal space directions. It is to note that if one vanishing point is at infinity, only the locus of the principal point may be found (the segment of the vanishing line defined by the other two points). If two vanishing points are at infinity, the third fixes the principal point (e.g. nadir point in vertical imagery) but the camera constant cannot be found. It is noted that the above constraints could be introduced into the algorithm to accept infinite vanishing points, too. Here, however, the primary intention is to perform full camera calibration from a single image. Consequently, in the search process only triplets of finite vanishing points forming acute triangles are considered (orthogonality criterion). Further, a principal point (orthocentre of the triangle) is generally expected to lie close to the image centre, while computed camera constants should have 'reasonable' values (camera criterion). Vanishing point triplets which satisfy these criteria are accepted and sorted according to their total vote; that with the highest score is chosen as the final triplet of dominant vanishing points.

As a result of this process it is in fact possible that irrelevant 3D lines whose projection planes happen to pass through a vanishing point may be mistakenly assigned to this vanishing point, or that image segments close to vanishing lines might be accepted to participate in the estimation of two vanishing points. Such misinterpretations in the positions of 3D lines appear as inevitable in a fully automated process, yet no significant effect on calibration results is expected if the angular threshold α_0 is kept small.

2.2. Camera calibration algorithm

The algorithm for camera calibration involves two sets of equations (linked to each other through common unknowns), namely those of line-fitting and those relating the coordinates of vanishing points to the parameters of camera geometry.

2.2.1. Line fitting

As mentioned, the estimation of camera constant and principal point location from three detected vanishing points of mutually orthogonal space directions is a straightforward task. Radial symmetric lens distortion may be estimated in a preceding step either through the 'plumb-line' technique (Brown, 1971; Devernay and Faugeras, 2001; Thormaehlen et al., 2003) based on the fact that a projected straight line should remain straight on the image or, on the other hand, along with vanishing point estimation (Becker and Bove, 1995; van den Heuvel, 1999; Bräuer-Burchardt and Voss, 2001) assuming that parallel straight lines in space are imaged as converging straight lines. Unlike approaches in which estimation of vanishing points precedes the calibration process, a main issue here is the line-fitting adjustment of image point observations for vanishing point estimation, interior orientation recovery and radial distortion compensation in one single step.

The observed image lines are constrained to converge to their corresponding vanishing point $V(x_V, y_V)$ according to following equation

$$(x - x_V)\cos\theta + (y - y_V)\sin\theta = 0 \qquad (2)$$

whereby x, y are the image coordinates of an individual point on a line. Introduction of the two coefficients k_1, k_2 of radial symmetric lens distortion into Eq. 2 results in:

$$\left[x - (x - x_o)(k_1 r^2 + k_2 r^4) - x_V \right]\cos\theta + \left[y - (y - y_o)(k_1 r^2 + k_2 r^4) - y_V \right]\sin\theta = 0 \qquad (3)$$

with (x_o, y_o) denoting the principal point and r the radial distance from it. This represents the observation equation for the common adjustment of all line segments (the root mean square distance of points x, y from the fitted line is minimized).

2.2.2. Vanishing points and interior orientation

2.2.2.1. The 'calibration sphere'

As mentioned already, two known vanishing points V_1, V_2 of orthogonal directions allow estimation of the camera constant only under the assumption of fixed principal point (usually the image centre). However, a constraint is always at work on the infinite possible locations of the projection centre O when only two vanishing points of orthogonal directions are known: O must view the segment $V_1 V_2$ under a right angle. Thus, in the 3D image space the geometrical locus of O is a sphere ('calibration sphere') of radius R, with the midpoint M of line segment $V_1 V_2$ as its centre and the distance $V_1 V_2$ as its diameter. For every possible projection centre O, the camera constant c equals its distance from the image plane, while the principal point P is its projection onto it (Fig. 2).

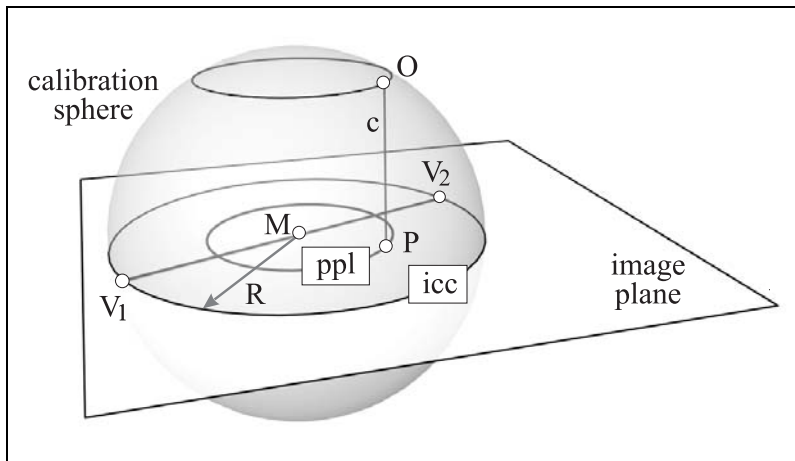Figure 2. Calibration sphere as locus of the projection centre O, principal point locus (*ppl*) for fixed camera constant c and isocentre circle (*icc*). V_1 and V_2 are the vanishing points of orthogonal directions, M is the midpoint of segment $V_1 V_2$.

The analytical equation of the sphere can be written as:

$$(x_o - x_m)^2 + (y_o - y_m)^2 + c^2 = R^2 \qquad (4)$$

whereby

$$x_m = \frac{x_1 + x_2}{2} \qquad y_m = \frac{y_1 + y_2}{2} \qquad R = \frac{1}{2}\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad (5)$$

with (x_1, y_1), (x_2, y_2) being the two vanishing points, (x_m, y_m) the centre of the sphere and R its radius. Each pair of vanishing points of orthogonal directions provides one such Eq. 4. Two pairs of such vanishing points (from the same image or from a different image sharing identical internal parameters) define a circle as the intersection of the corresponding spheres, whereas a third pair fixes the projection centre at the common point of the three 'calibration spheres' (Fig. 3). The projection centre O is, of course, that for which $c > 0$ (point 'above' the image plane), hence one should speak of 'calibration hemispheres' rather than spheres.
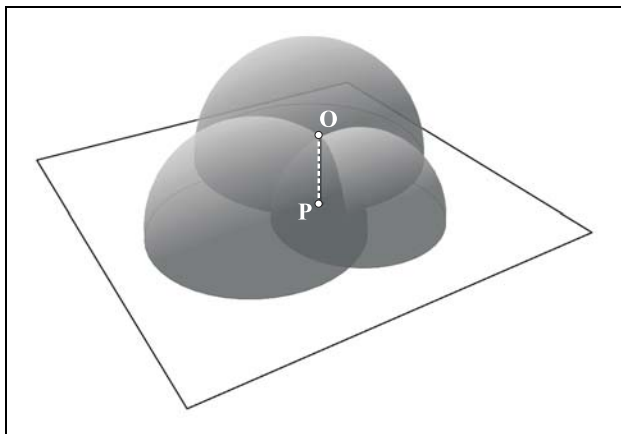

Figure 3. Projection centre O as intersection of three 'calibration spheres' and principal point P as its projection on the image plane.

It is worth noting that Eq. 4 can also be viewed as that of a circle on the image plane with centre (x_m, y_m) and radius $(R^2 - c^2)^{1/2}$. This circle *ppl* is the locus of the principal point P for a given camera constant c (Fig. 2). Indeed, c constrains the projection centre O on a circle of the sphere which is parallel to the image plane and, projected orthogonally onto it, defines the principal point locus (*ppl* in Fig. 2). It may also be observed that circle *icc*, in which the image plane intersects the sphere (Fig. 2), is in fact the locus of the 'isocentre', namely the point defined as the intersection of the image plane by the bisector of the angle formed by the image rays to the principal point and to the third vanishing point. This point is mentioned in photogrammetric literature mainly for its properties as regards measurement of angles on flat terrain (Church and Quinn, 1948).

Eq. 4 is the constraint relating vanishing points to the primary parameters of camera geometry and is used, along with Eq. 3, in the calibration adjustment which will be outlined in section 2.2.3. However, some additional remarks are useful here regarding image aspect ratio as well as the relation of the presented geometric interpretation to the image of the absolute conic and to plane-based calibration.

2.2.2.2.   Consideration of aspect ratio

Square image pixels have been assumed so far. For non-square pixels, the camera aspect ratio *a* expresses the relative scaling of image axes (for a normal camera $a = 1$). Let $a < 1$ be the scale factor for the vertical (y) coordinates of the image points which, thus scaled, are the orthogonal projections of the theoretical 'normal' points onto an image plane which is tilted by angle β ($\cos β = a$) about the x-axis (Fig. 4). The calibration sphere still holds for the normal image plane *nip* but the projection of the 'isocentre circle' *icc* is now an ellipse *ice* on the transformed (affine) image plane *aip*. The same is true for the locus of the principal point.
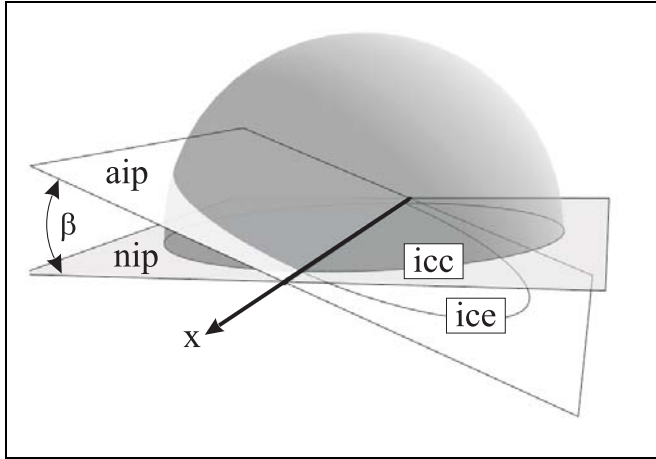

Figure 4. The isocentre ellipse (*ice*) is the orthogonal projection of the isocentre circle (*icc*) from the normal image plane (*nip*) onto the affine image plane (*aip*). The angle β between these two image planes with common x-axis is related to the aspect ratio *a* as $β = \cos^{-1} a$.

In this case, the calibration equation Eq. 4 becomes:

$$(x_o - x_m)^2 + \left(\frac{1}{a^2}\right)(y_o - y_m)^2 + c^2 = R^2 \tag{6}$$

with

$$R = \frac{1}{2}\sqrt{(x_1 - x_2)^2 + \left(\frac{1}{a^2}\right)(y_1 - y_2)^2}$$

2.2.2.3.   Calibration sphere and image of the absolute conic

In the field of computer vision, three vanishing points in orthogonal directions are also used in order to compute the image ω of the absolute conic, whose expression is subsequently decomposed to extract the camera internal parameters in the form of the 'calibration matrix' (Liebowitz et al., 1999; Sturm and Maybank, 1999). The conic ω is defined as:

$$\omega = \mathbf{K}^{-T}\mathbf{K}^{-1} \tag{7}$$

where **K** is the calibration matrix (Hartley and Zisserman, 2000). A pair of orthogonal vanishing points $\mathbf{v}_1$, $\mathbf{v}_2$ (in homogeneous representation) supplies a linear constraint on the entities of ω of the form

$$\mathbf{v}_1^T \omega \mathbf{v}_2 = 0 \tag{8}$$

Three such pairs allow estimation of the calibration matrix of a normal camera. However, with ω being an imaginary conic, the geometric relation of the vanishing points with the camera internal parameters is not obvious. Ignoring image aspect ratio and skewness, ω may be written as:

$$\omega = \begin{bmatrix} 1 & 0 & -x_o \\ 0 & 1 & -y_o \\ -x_o & -y_o & x_o^2 + y_o^2 + c^2 \end{bmatrix} \tag{9}$$

Introduction of Eq. 9 into Eq. 8, with $\mathbf{v}_1 = (x_1, y_1, 1)^T$, $\mathbf{v}_2 = (x_2, y_2, 1)^T$ in inhomogeneous representation, yields Eq. 4. Thus, the constraint of Eq. 8 and the equation of the 'calibration sphere' are equivalent for inhomogeneous notation of vanishing points.

2.2.2.4. *Calibration sphere and plane-based calibration*

The calibration sphere is also relevant in the case of plane-based calibration, a process which estimates interior orientation using homographies **H** between a plane with known Euclidean structure and its images (Zhang, 2000). The two basic equations of plane-based calibration are:

$$\mathbf{h}_1^T \omega \mathbf{h}_2 = 0 \tag{10}$$

$$\mathbf{h}_1^T \omega \mathbf{h}_1 = \mathbf{h}_2^T \omega \mathbf{h}_2 \tag{11}$$

with $\mathbf{h}_1 = [H_{11}, H_{21}, H_{31}]^T$, $\mathbf{h}_2 = [H_{12}, H_{22}, H_{32}]^T$ being the first two columns of the **H** matrix. Each homography supplies two such constraints on the elements of ω. Viewing the problem of plane-based calibration in a more intuitive geometric context, Gurdjos et al. (2002) proved that the solution is equivalent to intersecting circles ('centre circles'). Such a circle is the locus of the projection centre when space-to-image homography is known, and it may be obtained as the intersection of a sphere (centre sphere) and a plane (centre plane). In photogrammetric terminology, this plane (typically defined in aerial images as the vertical plane containing the camera axis) is the 'principal plane'. In fact, Eq. 10 is equivalent to the calibration sphere presented here. Eqs. (10) and (4) can be written as:

$$D = \mathbf{h}_1^{\mathrm{T}} \omega \mathbf{h}_2 = 0$$
$$G = (x_o - x_m)^2 + (y_o - y_m)^2 + c^2 - R^2 = 0$$

Taking into account that $\mathbf{v}_1 = [H_{11}, H_{21}, H_{31}]^{\mathrm{T}}$, $\mathbf{v}_2 = [H_{12}, H_{22}, H_{32}]^{\mathrm{T}}$ with respect to the elements of homography $\mathbf{H}$ and substituting them into G, it can be proved that G is equivalent to D on the condition that $H_{31}H_{32} \neq 0$. This indicates that the first equation of plane-based calibration (Eq. 10) is indeed that of the calibration sphere when both vanishing points are finite.

2.2.3. The camera calibration adjustment

Coming now to the camera calibration adjustment, this involves observation equations – one Eq. 3 for each point on every line segment – and the constraint equations Eq. 4 which introduce an additional unknown, the camera constant c. Common unknowns are the coordinates of all vanishing points as well as the coordinates x_o, y_o of the principal point (as said, no distinction is made here between the point of best symmetry of distortion and the principal point). If Eqs. (4) are treated as strict constraints, solution of the system of observation equations (Eqs. 3) requires approximate values for the coefficients of radial distortion. However, if Eqs. (4) are treated as observation equations with very large weights, it appears that initial values for k_1, k_2 may be set to zero, at least as experienced with the moderately wide-angle lenses with field of view up to 65° used in our experiments. It is also noted that the initial values for the location of vanishing points are supplied by the detection process of section 2.1.2, from which approximate values for the interior orientation parameters are computed through the equations given in Gracie (1968).

After the first iteration, the estimated coefficients of radial distortion are used to correct the initial (distorted) extracted edge points. The edge-linking process is repeated to allow formation of longer line segments using a reduced threshold of line 'straightness'. Thus, new candidacies for vanishing points emerge, and the search process results in the final triplet of vanishing points. Their position is then optimized with a least squares adjustment to provide refined values for line slopes. In order to identify convergent line segments which belong to the same line, all segments are checked in pairs. If the root mean square distance of all points of both segments from the fitted line equation of the other segment is below a threshold, such segments are seen as parts of the same image line and are treated as such in the next iterations.

In a previous publication, one single equation incorporating line-fitting, camera calibration and estimation of image attitude using vanishing points was proposed (Grammatikopoulos et al., 2003). Yet that approach appears to be more sensitive to image noise and outliers than the 'calibration hemispheres' which do not involve the image rotation matrix. Thus, the presented algorithm performs exclusively camera calibration (therefore, it is irrelevant which particular vanishing points are assigned to the X, Y or Z axes). Finally, it is reminded that, next to its applicability to single images with three finite vanishing points of mutually orthogonal directions, a basic advantage of this general algorithm is that further pairs of orthogonal vanishing points from the same or different views may be combined to estimate a common interior orientation. While three vanishing points of orthogonal directions allow finding only the primary internal camera parameters, further interior orientation parameters such as the image aspect ratio can be estimated at the expense of an additional pair of vanishing points from the same or a different image (Grammatikopoulos et al., 2004).

3.   Experimental results and discussion

The experimental evaluation with real data involves two groups of 800×600 images, taken with different cameras and shown in Fig. 5 (set I, four images) and Fig. 6 (set II, six images). These are typical terrestrial images, and thus allow to test the approach under commonly existing practical conditions, in the sense that they have been acquired with small upward tilts, i.e. they have rather weak perspectives in the vertical direction (vanishing point far from image centre). This geometry has a direct effect on the precision of the estimated x-coordinate (x_o) of the principal point (van den Heuvel, 1999). Besides, the two first images in set II have been taken also with an in-plane rotation of about 90°, which results in a similar effect on the y-coordinate of the principal point (y_o). In set I, the second and third images have been acquired from standpoints which are very close to each other.

Figure 5. Set I of images used in the evaluation.

Figure 6. Set II of images used in the evaluation.

The steps of line extraction and grouping are illustrated in Figs. 7 and 8. Initial extracted edges which exceeded 20 pixels in length are shown on the left. Using the angular threshold $\alpha_0 = 2°$, these edges were subsequently assigned to three dominant vanishing points (centre). The final line segments, which were generated after the correction of radial lens distortion and then used for calibration, are seen on the right. Two segments were treated as part of the same image line if the overall root mean square distance of their points from the line equation of the other segment did not exceed 2 pixels.
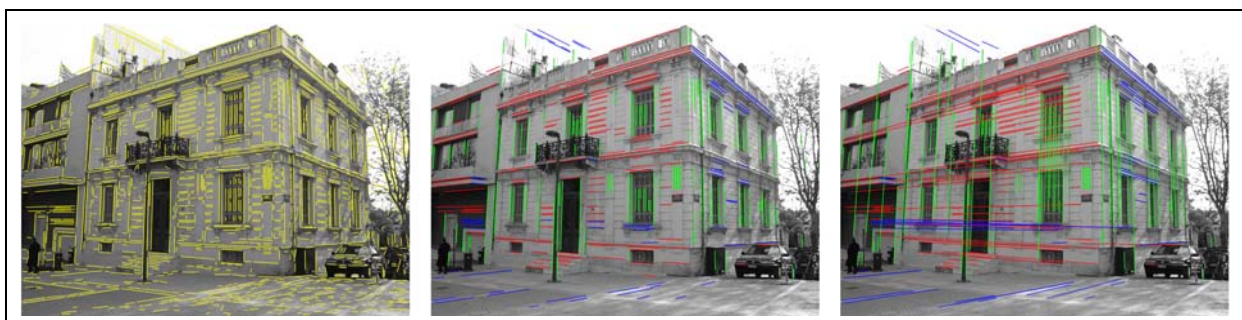

Figure 7. Initial extracted edges (left), line segments grouped in dominant vanishing points (centre) and final segments (right) after correction of lens distortion [third image in Fig. 5].

Figure 8. Initial extracted edges (left), line segments grouped in dominant vanishing points (centre) and final segments (right) after correction of lens distortion [last image in Fig. 6].

Tables 1 and 2 present the calibration results from vanishing points for each image as well as for the common adjustment of all four and six images of sets I and II, respectively, along with the estimated precision of parameter values, the standard error σ_o of the adjustment and the total number of participating lines and points. The rather large values of σ_o in Tables 1 and 2 are due to the threshold for segment linking (the initial extracted edges gave values for σ_o about half as large), which however is necessary if longer segments are to be formed. In order to obtain directly comparable values for the camera constant and the coefficients of the distortion polynomial, the latter must be brought to some 'calibrated' form. This was done here according to Karras et al. (1998) by minimizing through integration the sum of squares of distortion values up to the radial distance $r_{max} = 400$ pixels. Tables 3 and 4 present the three coefficients of the new polynomial $k_o r + k_1 r^3 + k_2 r^5$, along with the corresponding calibrated values $c_{cal} = c/(1+k_o)$ for the camera constant. The calibrated curves for both image sets are seen in Fig. 9.

Table 1. Camera calibration results from vanishing points
and bundle adjustments without control points (*b*) and with control points (*B*) – Set I

images	c (pix)	x_o (pix)	y_o (pix)	$k_1 (\times 10^{-7} pix^{-2})$	$k_2 (\times 10^{-13} pix^{-4})$	σ_o (pix)	# lines	# points
1	798.6 ± 0.9	-1.5 ± 2.0	-6.2 ± 0.8	-2.79 ± 0.12	5.57 ± 0.63	0.53	206	3877
2	790.4 ± 0.8	-1.9 ± 1.7	-4.1 ± 0.7	-2.34 ± 0.10	3.75 ± 0.47	0.59	231	4612
3	791.3 ± 0.7	-1.9 ± 1.6	-3.8 ± 0.6	-2.46 ± 0.09	4.32 ± 0.43	0.57	230	4674
4	798.2 ± 0.6	-5.1 ± 1.3	-6.8 ± 0.7	-2.61 ± 0.10	5.15 ± 0.51	0.58	215	5109
all	795.0 ± 0.4	-6.0 ± 0.7	-4.9 ± 0.3	-2.51 ± 0.05	4.56 ± 0.25	0.57	882	18272
b	*795.4 ± 1.6*	*0.8 ± 3.0*	*3.2 ± 5.9*	*-3.21 ± 0.17*	*4.11 ± 1.10*	*0.13*	*-*	*-*
B	*794.9 ± 1.0*	*4.4 ± 0.7*	*1.7 ± 0.6*	*-3.23 ± 0.12*	*3.86 ± 0.87*	*0.13*	*-*	*-*

Table 2. Camera calibration results from vanishing points
and bundle adjustment without control points (b) – Set II

images	c (pix)	x_o (pix)	y_o (pix)	$k_1 (\times 10^{-8} pix^{-2})$	$k_2 (\times 10^{-13} pix^{-4})$	σ_o (pix)	# lines	# points
1	896.2 ± 0.8	6.8 ± 0.6	-15.3 ± 0.9	-8.65 ± 0.59	-1.64 ± 0.25	0.45	124	2895
2	897.3 ± 0.8	11.4 ± 0.6	-10.9 ± 1.1	-9.03 ± 0.63	-2.38 ± 0.26	0.44	134	2826
3	908.2 ± 1.7	14.2 ± 2.1	-6.8 ± 1.4	-11.67 ± 0.97	-1.39 ± 0.42	0.40	88	1840
4	900.2 ± 1.2	13.9 ± 1.6	-8.3 ± 0.8	-4.30 ± 0.83	-3.97 ± 0.33	0.39	94	2138
5	902.5 ± 1.4	3.5 ± 1.8	-5.5 ± 1.1	-7.67 ± 0.80	-2.38 ± 0.29	0.36	84	1999
6	903.6 ± 1.3	18.5 ± 1.3	-2.6 ± 0.8	-7.47 ± 0.70	-3.41 ± 0.26	0.39	88	2248
all	899.2 ± 0.4	10.7 ± 0.4	-9.0 ± 0.3	-7.82 ± 0.29	-2.56 ± 0.12	0.42	612	13946
b	*898.2 ± 3.6*	*6.6 ± 1.6*	*-7.7 ± 1.6*	*-8.97 ± 1.24*	*-2.54 ± 0.52*	*0.27*	-	-

Table 3. Transformed distortion coefficients and
corresponding calibrated camera constant c_{cal} [= $c/(1+k_o)$] – Set I

images	c_{cal} (pix)	$k_o(\times 10^{-2})$	$k_1(\times 10^{-7}\,\mathrm{pix}^{-2})$	$k_2(\times 10^{-13}\,\mathrm{pix}^{-4})$
1	782.6	2.031	-2.733	5.457
2	776.4	1.807	-2.298	3.682
3	776.9	1.857	-2.414	4.240
4	783.3	1.908	-2.560	5.052
all	780.3	1.878	-2.463	4.474
b	*775.4*	*2.573*	*-3.127*	*4.004*
B	*774.6*	*2.618*	*-3.145*	*3.759*

Table 4. Transformed distortion coefficients and
corresponding calibrated camera constant c_{cal} [= $c/(1+k_o)$] – Set II

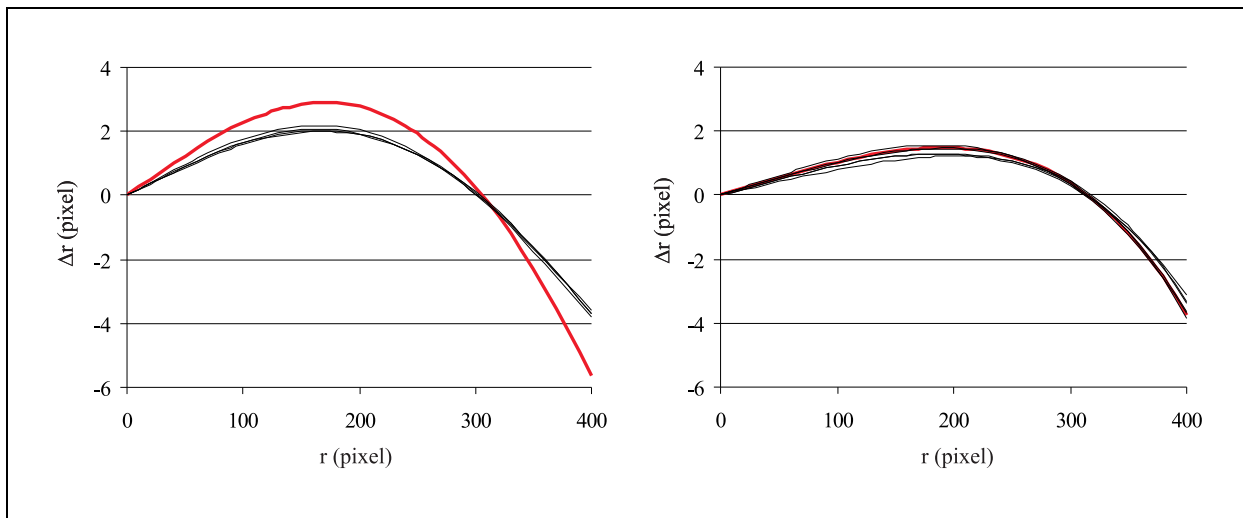| images	c_{cal} (pix)	$k_o(\times 10^{-2})$	$k_1(\times 10^{-8}\,\mathrm{pix}^{-2})$	$k_2(\times 10^{-13}\,\mathrm{pix}^{-4})$
1	887.3	1.002	-8.563	-1.624
2	887.4	1.118	-8.929	-2.353
3	896.9	1.254	-11.455	-1.373
4	892.7	0.843	-4.264	-3.936
5	893.7	0.990	-7.594	-2.356
6	893.9	1.083	-7.389	-3.373
all	890.0	1.024	-7.740	-2.534
b	*888.2*	*1.130*	*-8.867*	*-2.512*

Figure 9. Calibrated radial lens distortion curves (left: image set I; right: image set II) from vanishing points (four thin lines) and from multi-image bundle adjustment without control (thick red line).

Additionally, Tables 1 to 4 include corresponding results from self-calibrating bundle adjustments without use of ground control (b), which allows to compare estimations not based on external information. For set I, the block consisted of five convergent overlapping images of a different object with considerable extension in depth. Since here control points were indeed available, calibration results based on them are also included in Tables 1 and 3 (B). Note the significant uncertainty of the estimated principal point location in the absence of ground control. In the case of set II, the block consisted of eight overlapping images of the same 3D object, while no external control was at hand. Most points used in the first block

were signalized, which accounts for the smaller standard error of the adjustment as compared to that from the second block.

Calibration results from vanishing points of single images appear as being rather consistent. Referring to Tables 3 and 4 for the two cameras, all differences among individual values for the camera constant c as well as all differences of these individual c-values from the corresponding c-values estimated by bundle adjustment do not exceed 1.1%. The differences of the mean of individual c-values estimated through vanishing points from the c-value obtained with bundle adjustment with no control are 0.6% and 0.4% for the two cameras. The respective values for the difference of the c-value obtained by the common adjustment of all individual images from the bundle solution are 0.7% and 0.2%.

Regarding radial distortion, it is seen from Fig. 9 that the use of vanishing points yields almost identical curves for set I (maximum difference is 0.2 pixels). The curves for set II are also in agreement, with maximal differences at the image corner not exceeding 0.7 pixels; the curve from bundle adjustment passes among these curves. In image set I, however, results from vanishing points seem to significantly underestimate distortion compared to bundle adjustment, with maximal differences reaching 2 pixels at the image corners (obviously, this difference has an effect on the calibrated c-values of Table 3). This discrepancy regarding the particular images of set I might be attributed to the rather poor distribution of lines which do not cover adequately the whole image area. Nonetheless, one should not fail to mention an inherent limitation of this approach, namely that typical images are not likely to have both a strong perspective and lines recorded near the image corners.

As mentioned, the precision of the estimated principal point coordinates x_o, y_o is indeed lower in the direction perpendicular to that of the 'weak' vanishing point (y_o in the first two images of set II, x_o in all other cases). Furthermore, differences in the location of the principal point appear here as being large. However, differences of several pixels are generally to be expected, actually not only in cases based on vanishing points (e.g. compare results given in Wang and Tsai, 1990) but also among bundle adjustments of different image blocks. Ruiz et al. (2002) point to the fact that when standard calibration algorithms are applied to different image sequences from a fixed camera they generally display acceptable stability in the camera constant, but the principal point location is estimated with large variability. In this context, even if calibration data from bundle adjustment are at hand a direct comparison of parameter values may not be conclusive.

Indeed, Zhang et al. (1996) have demonstrated that it is very difficult to estimate precisely by calibration the coordinates of the principal point; yet, even large variations in the principal point location do not affect significantly the 3D reconstruction. Thus, the discussion here will also involve object reconstruction as an additional evaluation criterion. In a sense, this may be regarded as providing the missing 'ground truth', whose lack is a problem when quantifying the performance of non-metric calibration approaches on the basis of real images (Swaminathan and Nayar, 2000).

In order to assess the actual impact of the differences among calibration results, the 3D point coordinates reconstructed through self-calibrating bundle adjustment without control points were correctly scaled and used as the reference point set. The model points reconstructed through bundle adjustments, in which the calibration results from individual images based on vanishing points were alternately introduced as fixed values, were then compared with the reference point set. Consequently, differences in 3D reconstruction may be exclusively attributed to the differences in calibration data. The standard errors σ_{XYZ} of 3D similarity

transformations for all cases of both cameras are presented in Table 5. In order to be comparable, these values have been referred to image space through the mean image scale of the respective block. Since in the block of the first camera the point coordinates were known with high precision, the results from adjustments with both self-calibration and fixed calibration were also compared to the geodetic coordinates (*g* in Table 5). It is noted that the objects of both bundle blocks displayed considerable differences in depth, which is clearly important when comparing the effects of calibration data.

Table 5. Accuracy of 3D object reconstruction referred to models
from bundle adjustment with no control (b) and to geodetic control (g)

	Set I (*g*)	Set I (*b*)	Set II (*b*)
calibration data	σ_{XYZ} *(pixel)*	σ_{XYZ} *(pixel)*	σ_{XYZ} *(pixel)*
from image 1	0.4	0.4	0.6
from image 2	0.4	0.4	0.4
from image 3	0.4	0.4	0.8
from image 4	0.5	0.6	0.7
from image 5	-	-	0.8
from image 6	-	-	1.0
from all images	0.4	0.5	0.3
bundle (b)	0.2	-	-

Differences in calibration results are responsible for these discrepancies in 3D reconstruction. Even so, however, calibration data obtained automatically from vanishing points on single images with no geodetic control allowed to perform 3D measurements (in the case of the first camera in an altogether different scene) with a mean accuracy not worse than 1 pixel in object space. For the first camera this has also been verified against ground control. Thus, the presented tests indicate that the automatic single-image calibration approach through vanishing points, as reported here, shows good consistency and gives results which may definitely serve as initial values, but also as known calibration data for tasks of moderate accuracy requirements.

4.   Concluding remarks

In this paper the development and implementation of a camera calibration algorithm from single images with three vanishing points of orthogonal space directions is presented. A particular feature of the approach is that the camera constant, the principal point location as well as the two coefficients of radial symmetric lens distortion and the image vanishing points are all estimated simultaneously. This succeeds iteratively in a unified least-squares adjustment of all image points belonging to lines which intersect in the three dominant vanishing points of the scene. A further feature of this one-step algorithm is that it is equally suitable for the joint calibration of independent images with three and/or two vanishing points. Extraction of line segments and their association with particular vanishing points runs fully automatically, based on a modification of the approach of Rother (2000), while edges are linked anew after the first estimation of lens distortion. The evaluation against results from self-calibrating bundle adjustments indicated that this approach may, in principle, provide information on camera geometry which could be useful not only in terms of starting values but also, on several occasions, as fixed camera data. Clearly, further evidence from different

images and cameras, as well as experimentation with the involved thresholds, is necessary for a more thorough assessment of the method.

Of course, this approach is subject to all limitations present when working with vanishing points, namely its accuracy depends not only on image noise or possible deviations of lines from orthogonality and parallelism, but also on the distribution of image lines over the frame, their length, as well as the particular perspective distortion of the imaged scene. In this sense, an important further task would be to study, and possibly quantify in terms of a metric, the effect of the shape of the triangle defined by three vanishing points of orthogonal space directions on the precision of camera calibration.

5.   References

Antone, M. Teller, S., 2000. Automatic recovery of relative camera rotations for urban scenes. Proc. IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, June 13-15, vol. 2, pp. 282-289.

Barnard, S.T., 1983. Interpreting perspective images. Artificial Intelligence 21, 435-462.

Bauer, J., Klaus, A., Karner, K., Zach, C., Schindler, K., 2002. MetropoGIS: a feature based city modeling system. Proc. Photogrammetric Computer Vision, ISPRS Comission III Symposium, September 9-13, Graz, Austria (in CDROM).

Becker, S. Bove, M., 1995. Semiautomatic 3-d model extraction from uncalibrated 2-d camera views. Proc. SPIE Visual Data Exploration and Analysis II, San Jose, California, 8-10 February, vol. 2410, pp. 447-461.

Bräuer–Burchardt, C., Voss, K., 2001. Image rectification for reconstruction of destroyed buildings using single views. In: B. Fisher, K. Dawson-Howe, C. O'Sullivan (eds.), Virtual and Augmented Architecture, Springer, Berlin, pp. 159-170.

Brillault-O'Mahony, B., 1991. New method for vanishing point detection. Computer Vision, Graphics, and Image Processing 54(2), 289-300.

Brown, D.C., 1971. Analytical calibration of close-range cameras. Photogrammetric Engineering 37(8), 855-866.

Caprile, B., Torre, V., 1990. Using vanishing points for camera calibration. International Journal of Computer Vision 4(2), 127-140.

Cantoni, V., Lombardi, L., Porta, M., Sicard, N., 2001. Vanishing point detection: representation analysis and new approaches. Proc.11th International Conference on Image Analysis and Processing, 26-28 September, Palermo, Italy, pp. 90-94.

Church, E., Quinn, A.O., 1948. Elements of Photogrammetry. Syracuse University Press, Syracuse, New York.

Cipolla, R., Drummond, T., Robertson, D.P., 1999. Camera calibration from vanishing points in images of architectural scenes. Proc. British Machine Vision Conference, 13-16 September, Nottingham, UK, pp. 382-391.

Devernay, F., Faugeras, O., 2001. Straight lines have to be straight. Machine Vision and Applications 13(1), 14-24.

Gallagher A. C., 2002. A ground truth based vanishing point detection algorithm. Pattern Recognition 35(7), 1527-1543.

Gurdjos, P., Crouzil, A., Payrissat, R., 2002. Another way of looking at plane-based calibration: the centre circle constraint. Proc. European Conference on Computer Vision, Copenhagen, May 28-31, Lecture Notes in Computer Science, Springer, Berlin, vol. 2353, pp. 252-266.

Gracie, G., 1968. Analytical photogrammetry applied to single terrestrial photograph mensuration. Proc. XI Congress of the International Society for Photogrammetry, Commission V, Lausanne, 8-20 July (unpaginated).

Grammatikopoulos, L., Karras, G., Petsa, E., 2002. Geometric information from single uncalibrated images of roads. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part B5), 21-26.

Grammatikopoulos, L., Karras, G., Petsa, E., 2003. Camera calibration approaches using single images of man-made objects. Proc. XIX CIPA International Symposium, Antalya, Turkey, 30 September - 4 October, pp. 328-332.

Grammatikopoulos, L., Karras, G., Petsa, E., 2004. Camera calibration combining images with two vanishing points. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 35 (Part 5), 99-104.

Hartley, R., Zisserman, A., 2000. Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge, UK.

Karras, G., Patias, P., Petsa, E. 1993. Experiences with digital rectification on non-metric images when ground control is not available. Proc. XV CIPA International Symposium, Bucharest (unpaginated).

Karras, G., Mountrakis, G., Patias, P., Petsa, E., 1998. Modeling distortion of super-wide-angle lenses for architectural and archaeological applications. International Archives of Photogrammetry and Remote Sensing 32 (Part 5), 570-573.

Kosecká, J., Zhang, W., 2002. Efficient computation of vanishing points. Proc. IEEE International Conference on Robotics and Automation, 11-15 May, Washington, DC, USA, pp. 223-228.

Liebowitz, D., Criminisi, A., Zisserman A., 1999. Creating architectural models from images. Eurographics '99, Computer Graphics Forum 18(3), 39-50.

Lutton, E., Maitre, H., Lopez-Krahe, J., 1994. Contribution to the Determination of Vanishing Points Using Hough Transform. IEEE Transactions on Pattern Analysis and Machine Intelligence 16(4), 430-438.

Magee M., Aggarwal J., 1984. Determining vanishing points from perspective images. Computer Vision, Graphics, and Image Processimg 26, 256–267.

Merritt, E.L., 1958. Analytical Photogrammetry. Pitman Publishing Co., New York,

Quan, L., Mohr, R., 1989. Determining perspective structures using hierarchical Hough transform. Pattern Recognition Letters 9, 279-286.

Rasmussen, C., 2004. Texture-based vanishing point voting for road shape estimation. Proc. British Machine Vision Conference, 7-9 September, London, UK (http://www.bmva.ac.uk/bmvc/2004/papers/paper_261.pdf, last access date: February 17, 2007).

Rother, C., 2000. A new approach for vanishing point detection in architectural environments. Proc. British Machine Vision Conference, 11-14 September, Bristol, UK, pp. 382–391.

Rother, C., 2002. A new approach to vanishing point detection in architectural environments. Image and Vision Computing 20(9-10), 647-655.

Ruiz A., López-de-Teruel P.E., García-Mateos G., 2002. A note on principal point estimability. Proc. 16th International Conference on Pattern Recognition (ICPR '02), 11-15 August, Quebec, Canada, vol. 2, pp. 304- 307.

Shufelt, J.A., 1999. Performance evaluation and analysis of vanishing point detection techniques. IEEE Transactions on Pattern Analysis and Machine Intelligence 21(3), 282–288.

Straforini, M., Coelho, C., Campani M., 1993. Extraction of vanishing points from images of indoor and outdoor scenes. Image and Vision Computing 11(2), 91-99.

Sturm, P., Maybank, S. J., 1999. A method for interactive 3D reconstruction of piecewise planar objects from single images. Proc. British Machine Vision Conference, 13-16 September, Nottingham, UK, pp. 265-274.

Swaminathan R., Nayar, S., 2000. Non-metric calibration of wide-angle lenses and polycameras. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(10), 1172-1178.

Thormaehlen, T., Broszio, H., Wassermann, I., 2003. Robust line-based calibration of lens distortion from a single view. Proc. Mirage 2003, Rocquencourt, France, 10-12 March, pp. 105-112.

van den Heuvel, F. A., 1998. Vanishing point detection for architectural photogrammetry. International Archives of Photogrammetry and Remote Sensing 32 (Part 5), 652-659.

van den Heuvel, F., 1999. Estimation of interior orientation parameters from constraints on line measurements in a single image. International Archives of Photogrammetry and Remote Sensing 32 (Part 5W11), 81-88.

Wang L.-L., Tsai W.-H., 1990. Computing camera parameters using vanishing-line information from a rectangular parallelepiped. Machine Vision and Applications 3, 129-141.

Wenzel, F., Grigat, R.-R., 2006. Representing directions for Hough Transforms. Proc. International Conference on Computer Vision Theory and Applications, 25-28 February, Setúbal, Portugal, vol. 2, pp. 116–121.

Zhang Z., Luong Q.-T., Faugeras, O., 1996. Motion of an uncalibrated stereo rig: self-calibration and metric reconstruction. IEEE Transactions on Robotics and Automation 12(1), 103-113.

Zhang Z., 2000. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11), 1330-1344.