RECOVERING THE VANISHING SELF-POLAR TRIANGLE FROM A SINGLE VIEW OF A PLANAR PATTERN

Pierre Gurdjos, René Payrissat

IRIT-UPS, 118 route de Narbonne
31062 Toulouse Cedex, France
{Pierre.Gurdjos,Rene.Payrissat}@irit.fr

ABSTRACT

We are concerned with the recovery of the metric structure of the 3D space from the metric structure of a plane, given as the world-plane to image homography. Recent works have pointed out that the metric structure of the 3D space can be defined by a set of three mutually conjugated (vanishing) points with respect to the imaged absolute conic (the vanishing self-polar triangle) and three additional factors, defined up to a scale factor. Given that the homography H only allows the image to inherit two of these points and two of these factors, we show how the (three) unknown quantities can be recovered from H, by investigating some camera constraints which are discussed in the paper. We propose two solutions in the cases of (i) of known principal point coordinates (ii) of zero-skew and known pixel aspect ratio. We demonstrate that, under these assumptions, the third vanishing point, associated with the direction orthogonal to the reference plane, lies on a line that we call central line. Adding an additional constraint, zero-skew for (i) and known camera height for (ii), a direct solution is found. Results on real images have proved to be quite accurate and encouraging for the use of our approach.

1. INTRODUCTION

In the image plane, the vanishing points are the perspective projections of three orthogonal direction vectors (considered as points at infinity) with respect to a certain 3D world coordinate system. The vanishing self-polar triangle is a triangle with vanishing points as vertices that is self-polar, which means that the three (vanishing) points are mutually conjugated with respect to the absolute conic. Fundamentally, the absolute conic [1] is the key object for (self-)calibration and its image, represented by a 3×3 symmetric matrix ω with five degrees of freedom, carries the metric information of the projective space. Consequently, the recovery of the vanishing self-polar triangle is closely connected to the problem of recovering the metric structure of the scene and it has been shown[2] [3] that a *natural* camera calibration (*ie* with a square-pixel with zero-skew constraint) can be done, in a single view, from the vanishing self-polar triangle. In this paper, we are concerned with the recovery of the vanishing self-polar triangle from the single view of a planar pattern, given the related world-plane to image homography. It is obviously impossible without additional constraints about camera or scene. In previous works, at least two perpendicular planar patterns[3] or three images of a single planar pattern were required [4]. This calibration problem from a planar pattern recently arouse interest in computer vision [5] because, in many man-made environments, planes are widely present and easily identifiable. Given a single view, a plane-to-plane mapping (*ie* the world plane to image homography) can be estimated and a lot of highly interesting techniques have been recently developped for this purpose [6] [7]. While the image of the absolute conic can be given by a set of three mutually conjugated (vanishing) points [3] $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ associated with three additional factors $\alpha : \beta : \gamma$, the homography only allows the image to inherit two of these points, say $\mathbf{v}_1$,$\mathbf{v}_2$, and two of these factors, say $\alpha : \gamma$ (equivalently, an homography only provides two constraints on the matrix ω, as it is described in [4]). We will investigate the camera and scene constraints under which, assuming that $\mathbf{H}$ has been estimated, the recovery of $\mathbf{v}_3$ and β, *ie* that fully defines matrix ω, can be carried out. In fact we point out that our problem is equivalent to the problem of recovering the parametrization $[\mathbf{n}, d]$ of the plane associated with the planar pattern, where $\mathbf{n}$ is the normal vector and d the plane offset. The paper is constructed as follows. After giving some geometric background, we demonstrate that, under the assumptions of known principal point coordinates or squared pixels with zero-skew, the third vanishing point lies on a line that we call central line. A direct solution is found with an additional constraint, about camera height, relevant to application we dealt with. We end up the paper with applications to real images.

2. PROBLEM STATEMENT AND GEOMETRIC BACKGROUND

In this paper, a 2D coordinate vector is denoted by $\mathbf{x}$ while a homogeneous (augmented by adding 1) vector is denoted by $\tilde{\mathbf{x}} = [\mathbf{x}, 1]^T$. The notation $\alpha : \beta : \gamma$ means three scalars defined up to a scale factor. The coordinates of a 3D world point $\mathtt{W}=[X, Y, Z]^T$ and its image coordinates $\mathbf{m} = [u, v]^T$ are related by the equation $s\tilde{\mathbf{m}} = \mathbf{P}[\mathtt{W}, 1]^T$ where $s \neq 0$ is an arbitrary scale factor and $\mathbf{P}$ is the 3×4 perspective projection matrix. $\mathbf{P}$ is usually decomposed as $\mathbf{P} = \mathbf{A}[\mathbf{R} \mid \mathbf{t}]$ where $\mathbf{A}$ is the upper triangular matrix related to the internal parameters of the camera, *ie* the principal point coordinates (u_0, v_0), the scale factors κ_u, κ_v, the skewness factor s and $(\mathbf{R}, \mathbf{t})$ is a 3D rigid displacement. Without loss of generality, we assume that the reference plane, denoted π, has equation $Z = 0$ in the world frame. Consequently, the π-plane to image homography is the restriction of $\mathbf{P}$ to π and can be represented by any matrix $\mathbf{H} = \lambda \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_4 \end{bmatrix}$, $\lambda \neq 0$, where $\mathbf{p}_i$ is the i-th column vector of $\mathbf{P}$. It is easy to verify that the reference plane π is parametrized, in the camera frame, by the vector $[\mathbf{n}, d] = [\mathbf{A}^{-1}\mathbf{p}_3, \mathbf{p}_3^T\mathbf{A}^{-T}\mathbf{A}^{-1}\mathbf{p}_4]$ where symmetric matrix $\mathbf{A}^{-T}\mathbf{A}^{-1}$ is the usual representation of the imaged absolute conic. We introduce now a slightly different decomposition of ω that will be very useful for our problem. The image of the absolute conic can be given by a set of three mutually conjugated (vanishing) points [3] $\mathbf{v}_1$,$\mathbf{v}_2$,$\mathbf{v}_3$ associated with three additional factors $\alpha : \beta : \gamma$ (with only two independent). Indeed it can be shown that ω can be decomposed as:

$$\omega = \mathbf{V}^{-T}\mathbf{D}^{-1}\mathbf{V}^{-1} \tag{1}$$

where $\mathbf{V}$ is compounded of column vectors $\tilde{\mathbf{v}}_i = \mathbf{p}_i/p_{3i} = [u_i, v_i, 1]^T = [\mathbf{v}_i, 1]^T$ represent augmented coordinates of the three vanishing points imaging the direction vectors of the world coordinate system, and $\mathbf{D} = diag(\gamma, \alpha, \beta)$ is a diagonal matrix. The homography matrix $\mathbf{H}$ only allows the image to inherit two of these points, $\mathbf{v}_1$,$\mathbf{v}_2$ as well as the factor α (considering $\gamma = 1$). Assuming that matrix $\mathbf{H}$ can been estimated, our problem can be stated as the problem of recovering $\mathbf{v}_3$ and β, in order to compute $\mathbf{p}_3 = \pm\sqrt{\beta}\mathbf{v}_3$. We omit the proof but we can claim that $\sqrt{\beta}$ is always defined as a real number ; the choice of its sign depending whether the camera frame is direct or not. When it is done, the projection matrix P can also be recovered. We now detail how we have demonstrated that, under certain constraints about the camera model, the vanishing point $\mathbf{v}_3$ lies on a straight line (crossing the principal point) that we will define. Let ω_{ij} be the (i, j)-th element of the symmetric matrix ω.

Zero-skew constraint. It can be shown that the assumption of a rectangular image coordinate system is equivalent to write:

$$\omega_{12} = 0 \tag{2}$$

Principal point constraint. When the principal point is known, a simple change of image coordinates leads us to state $(u_0, v_0) = (0, 0)$. It can be shown that this substitution is equivalent to two constraints on the matrix ω that are:

$$\begin{cases} \omega_{13} = 0 \\ \omega_{23} = 0 \end{cases} \tag{3}$$

By solving system (3) from (1) with $\gamma = 1$ for unknowns u_3, v_3, β, we obtain the equation $L_1 u + L_2 v = 0$ of a straight line (crossing the principal point), called *central line* $\mathcal{L}$, where

$$L_1 = (\alpha v_2 + v_1) \qquad L_2 = -(\alpha u_2 + u_1)$$

which holds for the vanishing point coordinates $(u = u_3, v = v_3)$. On the other hand, the solving simultaneously yields the solution $\beta = -(\alpha v_2 + v_1)/v_3$.

It is worthy of note that adding a zero constraint to system (3) is equivalent to add equation (2) yielding the direct solution for

$$\mathbf{v}_3 = \begin{bmatrix} u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} (\alpha v_2 u_2 + u_1 v_1)/(\alpha u_2 + u_1) \\ (\alpha v_2 u_2 + u_1 v_1)/(v_1 + \alpha v_2) \end{bmatrix}$$

and for $\beta = -\frac{(\alpha^2 v_2 u_2 + \alpha u_1 v_2 + \alpha u_2 v_1 + u_1 v_1)}{(a v_2 u_2 + u_1 v_1)}$.

Square-pixel constraint. In practice, when (u_0, v_0) are unknown, we might always consider the principal point as the image center. Nevertheless it does not seem a reasonable assumption as soon as the image has been cropped from a larger image. In fact, if the camera model has a zero-skew $s = 0$ and squared pixels $\kappa_u = \kappa_v$ (or equivalently known aspect ratio), we can write:

$$\begin{cases} \omega_{12} &= 0 \\ \omega_{11} &= \omega_{22} \end{cases} \tag{4}$$

Indeed, with respect to an image orthonormal coordinate system, circular points[8] have image complex projective coordinates $[1, \pm i, 0]^T$ and, since they lie on the imaged absolute conic, we have (4). Like we did for the known principal point constraints, we solve system (4) with $\gamma = 1$ for unknowns u_3, v_3, β, and we obtain the *central line* $\mathcal{L}$ equation $L_1 u + L_2 v + L_3 = 0$ which holds for the vanishing point $(u = u_3, v = v_3)$, with:

$$\begin{aligned} L_1 &= (\alpha + 1)(u_2 - u_1) \\ L_2 &= (\alpha + 1)(v_2 - v_1) \\ L_3 &= -u_2^2 - v_2^2 + \alpha(u_1^2 + v_1^2) + (1 - \alpha)(u_1 u_2 + v_1 v_2) \end{aligned} \tag{5}$$

In this case, it may be easily verified that (i) the principal point of the image plane lies on $\mathcal{L}$, (ii) the property exhibited by Caprile and Torre in [2], saying that the principal point is the orthocentre of the triangle with vanishing points as vertices, only holds if the image coordinate system is orthonormal, *ie* with a unit aspect ratio and a zero skew. On the other hand, the solution of the system (4) simultaneously yields $\beta = \frac{(u_1 - u_3)^2(1+\alpha)}{\alpha(v_1 - v_2)^2 + \alpha^{-1}(v_2 - v_3)^2 + \xi}$ where $\xi = (2v_2^2 - 2v_1 v_2 - 2v_2 v_3 + 2v_1 v_3 - u_3^2 - u_1^2 + 2u_1 u_3)$. The case in which is β not defined can be used as a constraint on the domain of value of v_3. In two solutions given above (known principal point and square-pixel cosntraints), we have only two independent constraints while three are required, one ambiguity remains. We give an example how to solve it by adding an assumption about the camera external parameters.

Known camera height. The vanishing point $\mathbf{v}_3$ with coordinates (u_3, v_3) lies at intersection of the central line $\mathcal{L}$ with the vanishing line associated with any plane perpendicular to π. Let π^* be a plane perpendicular to π, defined in the image by its intersection with π. It can be verified that the vanishing line associated with π^* is represented by the vector $\mathbf{H}^{-T}\mathbf{l}$ where $\mathbf{l} = [0, -\frac{1}{d}, 1]^T$ are Plücker coordinates of a line involving the distance d of camera centre C to π^*. Practically, if π^* is the "ground" plane, only the height (*ie* the Z-coordinate) of the camera has to be known. This particularly applies in the case of snapshots taken by persons if their heights are known. Nevertheless, we showed that, when this information is not available, a partial knowledge of the camera orientation is equivalent. It includes the special case where the horizontal axis of the camera is "almost" parallel to the ground[9], which is very likely regarding the camera orientation in some applications (TV broadcast sport images for instance).

3. EXPERIMENTS

We applied our technique for recovering the direction orthogonal to a reference plane π (in the form of the vanishing point coordinates $\mathbf{v}_3$). Let π^* be the horizontal plane perpendicular to π like in Fig.1. The camera is approximately 1m away from π and 0.65m away from π^* while the image dimension is 512×512. Let $\mathbf{H}$, respectively $\mathbf{H}^*$, be the world-plane to image homography matrix that maps points in π, respectively π^*, onto points in the image. We have preliminarily computed $\mathbf{H}$ and $\mathbf{H}^*$ but *only* matrix $\mathbf{H}$ has been used as the input data of our algorithm ; $\mathbf{H}^*$ must be considered as verification data and remained unknown to our algorithm. The recovering of the vanishing point coordinates $\mathbf{v}_3$ with β have been implemented in two ways, which are illustrated in Fig.1. The former (a) corresponds to a zero-skew and principal point constraints and the latter (b) to square-pixel and kown camera height constraints. In (b), $\mathbf{v}_3$ has been computed as the intersection of the vanishing line of π^*, computed from the known camera height, with the central line $\mathcal{L}$ derived from the simplified camera constraint. We show that it is possible to insert virtual 3D objects in the image once $\mathbf{v}_3$ and associated factor β have been recovered, like the parallelepiped adjacent to both planes π and π^* (the construction of projection matrix $\mathbf{P}$ is straightforward). Indeed it is straightforward to then compute the projection matrix $\mathbf{P}$. In Fig.2, we present two examples where the 3D coordinate system of a reference plane is recovered under the principal point constraint and attached to an arbitrary point of the reference plane. The figure shows up the normal vector to the plane.

4. CONCLUSION

Recent works [2][6][3] have pointed out that the metric structure of 3D space can be defined by a set of three mutually conjugated (vanishing) points with respect to the imaged absolute conic and two additional factors. We show how these quantities can be recovered from a single plane to image homography, under some camera constraints which are discussed in detail. From a practical point of view, it indeed seems very interesting to exploit the metric properties of a plane in order to deduce metric properties of a second one, perpendicular to the first for instance.

5. REFERENCES

[1] O. Faugeras, *Three-dimensional Computer Vision : a geometric viewpoint*, MIT Press, Boston, USA, 1993.

[2] B. Caprile and V. Torre, "Using vanishing points for camera calibration," *IJCV*, vol. 4, no. 2, pp. 127–140, March 1990.

[3] D. Liebowitz and A. Zisserman, "Combining scene and auto-calibration constraints," in *Proc. 7th International Conference on Computer Vision, Kerkyra, Greece*, September 1999.

[4] Z. Zhang, "A flexible new technique for camera calibration," Tech. Rep. MSR-TR-98-71, Microsoft corporation, Dec. 1998 (updated, 1999).

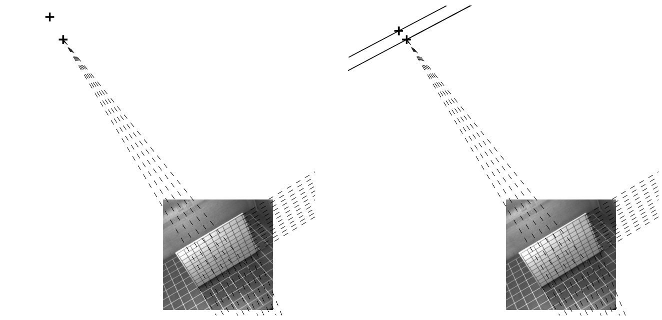
[5] B. Triggs, "Autocalibration from planar scenes," in *Proceedings of the 5th European Conference on Computer Vision (ECCV'98)*, Freiburg, Germany, June 1998, pp. 89–105.

[6] A. Criminisi, I. Reid, and A. Zisserman, "Single view metrology," in *Proc. 7th International Conference on Computer Vision, Kerkyra, Greece*, September 1999, pp. 434–442.

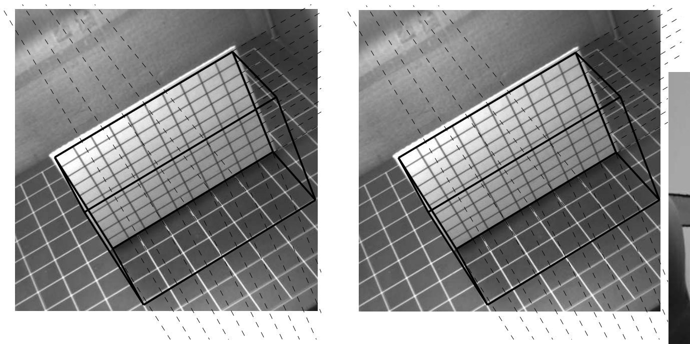
[7] D. Liebowitz and A. Zisserman, "Metric rectification for perspective images of planes," in *Proceedings of the Conference on Computer Vision and Pattern Recognition*, June 1998, pp. 482–488.

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

[9] P. Gurdjos and R. Payrissat, "About conditions for recovering the metric structures of perpendicular planes from the single ground plane to image homography," in *ICPR'2000 (to be published)*, Barcelona, Spain, September 2000.



(a-1) (b-1)



(a-2) (b-2)

Fig. 1. The four figures represent the same image of a scene consisting of a vertical planar patch π (materialized by a black grid with white background) put on an (horizontal) ground plane π^* (materialized by a white grid with dark grey background). Given the π-plane to image homography matrix $\mathbf{H}$, we suggest two ways of recovering the vanishing point $\mathbf{v}_3$ imaging a direction vector orthogonal to π. In the former, we only use the fact that the principal point is supposed to be the centre of the image (=256,256) while in the latter, we assume to deal with a simplified model of camera, *ie* with known aspect ratio (=1), and known camera height ($\simeq$0.65m). In figures (a-1) and (b-1), the plain cross corresponds to our estimation of $\mathbf{v}_3$. In order to compare, its awaited position and associated line pencil have been computed from $\mathbf{H}^*$ and drawn in dotted. In second place, once $\mathbf{v}_3$ and associated scale factor β recovered, we have also been able to recover the projection matrix $\mathbf{P}$. In (a-2) and (b-2), we show that it is henceforth possible to insert virtual 3D objects in the image, like the parallelepiped adjacent to both planes π and π^*. The way the parallelepiped is layed out on the ground plan shows off that metric properties has also been recovered.

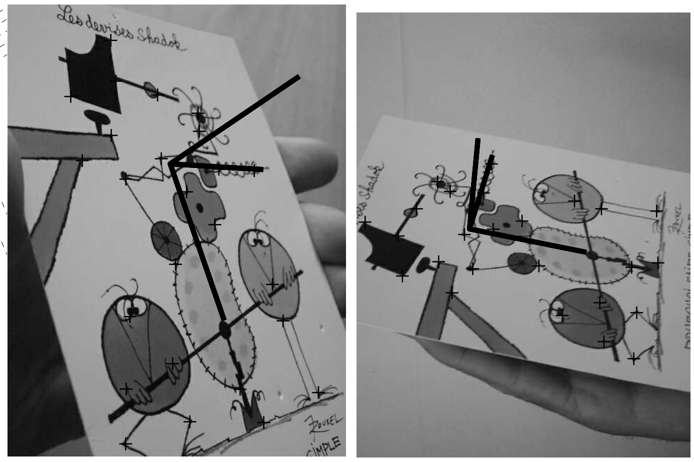


Fig. 2. Recovery of the 3D orientation of a reference plane from the single world-plane to image homography under the principal point constraint. The homography has been estimated from points represented with a white cross.