Infinite Homography Estimation Using Two Arbitrary Planar Rectangles

Jun-Sik Kim and In So Kweon

Dept. of EECS, KAIST,
373-1 Kusong-dong, Yuseong-Gu, Daejeon, Korea
jskim@rcv.kaist.ac.kr, iskweon@kaist.ac.kr
http://rcv.kaist.ac.kr

Abstract. In this paper, we propose a new method to estimate an infinite homography between two views containing two arbitrary planar rectangles. The proposed method does not require metric measurements, such as rectangle lengths or aspect ratios of the rectangles. We introduce the concept of semi-metric cameras and show that the semi-metric cameras derived from different views that see an identical 3D rectangle, can be regarded purely translating cameras whose pixel is zero-skewed. New parameterization for infinite homography is developed based upon the semi-metric space, and this parameterization is used to propose a new algorithm to estimate infinite homography. As a direct application, we apply our algorithm to autocalibration for a scene only with a few feature points on each rectangles.

1 Introduction

In the real world, there are many objects with two-dimensional planes and rectangular shapes, especially in outdoor urban environment. Cameras generally use planar CCD or CMOS type sensors. Therefore, the imaging process of planar objects can be described as a 2D to 2D transformation [1]. Furthermore, in the case of multiple views, the transformation between imaged planes can be also considered 2D to 2D, and is called *plane induced homography*. Plane induced homography offers a useful tool to describe scenes with planar objects from two or more views, as shown in plane + parallax approaches [1, 2, 3, 4, 5].

Among the plane induced homographies, a particularly important one is an *infinite homography*. The infinite homography is the homography induced by the plane at infinity with some important properties. First, it maps features on the plane at infinity of one view, such as vanishing points, vanishing lines and images of absolute conic, to another views. Second, it can be used to find affine and metric reconstructions from projective ones. This means we can calibrate a camera from image sequences using the infinite homographies. Additionally, we can reduce the search region for stereo matching through mapping with the infinite homography. Detailed explanations about these issues can be found in [1]. Note that the infinite homography between two views depends only on the rotation between the cameras capturing the views and the intrinsic parameters of them.

P.J. Narayanan et al. (Eds.): ACCV 2006, LNCS 3852, pp. 1–10, 2006.

Three methods are commonly used to estimate the infinite homography between two views. The first method uses camera motion constraints. If we use a purely rotating camera to capture the images, the homography induced by any plane on the image is the infinite homography. Although this method is easy to apply, it requires the use of rotating cameras. The second method uses strong scene constraints. If we have three vanishing points in each view with a fundamental matrix, the infinite homography can be estimated. Similarly it can be calculated from corresponding vanishing lines and vanishing points with the fundamental matrix. This requires the identification of three vanishing points and vanishing lines, however it may be difficult to find the features in infinity. The third method is a stratified approach. Once we find an affine reconstruction and projectively transformed plane at infinity, we can find the infinite homography from the projective projection matrix and the plane at infinity. The most difficult part of this approach is to build an affine reconstruction from the projective one. It requires some constraints of the scene and the camera, or the modulus constraints [6] for a static camera.

In this paper, we propose a new method to linearly estimate the infinite homography from images containing two arbitrary rectangles. The term "arbitrary" implies that we do not have information regarding the lengths, the aspect ratios, and the relative poses of the two rectangles. This method uses information about the parallelism and orthogonality, however this method does not require finding the vanishing points or the vanishing lines explicitly, which can be difficult for some rectangles. Furthermore, estimating epipolar geometry is also not required to estimate the infinite homography. Only tracking two rectangles between two views is needed.

In Sect. 2, we introduce the concept of semi-metric cameras and discuss some of their properties, such as image of absolute conic and special form of camera matrix. Sect. 3 discusses ways to parameterize the infinite homography using semi-metric cameras and to estimate the infinite homography using the proposed parameterization with two imaged rectangles. In Sect. 4, we show an important application of the infinite homography - the autocalibration of cameras - using the proposed algorithm. We conclude this paper in Sect. 5.

2 Semi-metric Cameras

We have introduced the concept of semi-metric space, defined as the sub-space of affine space [7]. In semi-metric space, orthogonal features are preserved, however the aspect ratio between two orthogonal axes is not preserved.

Assuming that there is a rectangle with an unknown aspect ratio in 3D space and a view capturing the rectangle in a general position, we can find a homography to make the projectively distorted rectangle to align the orthogonal axis of the rectangle. The warped image is called as semi-metric image. To make semi-metric images, two methods are used [7].

The first method uses vanishing points whose directions are orthogonal to each other. Warping the vanishing points to infinite points makes a semi-metric image with warping matrix defined as

$$\mathsf{H}_{sm} = \left[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{x}_c\right]^{-1}$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are vanishing points orthogonal to each other and $\mathbf{x}_c$ is an arbitrary point, as shown in Fig. 1.

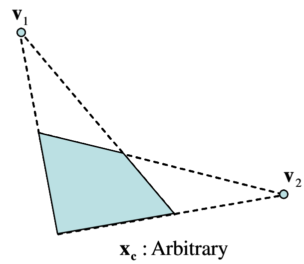


Fig. 1. Elements of semi-metric transformation matrix from vanishing points

Warping from the projected rectangle to a standard predefined rectangle with a known aspect ratio is sufficient to warp to a semi-metric image. Fig. 2 shows the concept of the warping method using a standard rectangle. Note that an aspect ratio of the warped rectangle can be set arbitrarily. For example, in Fig. 2, the aspect ratio is set to one.

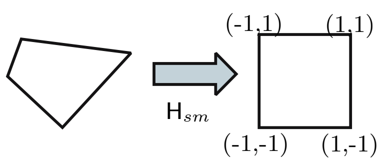


Fig. 2. Semi-metric warping using a standard rectangle

With a semi-metric image, the following theorem can be proven [7].

Theorem 1. *In semi-metric space, the ICDCP is given as* $\mathrm{diag}\left(R_m^2, R_{sm}^2, 0\right)$ *where* R_m *is the aspect ratio of the model rectangle, and* R_{sm} *is the aspect ratio of a semi-metric warped rectangle.*

Because the ICDCP in semi-metric space is expressed as $\mathrm{diag}(R_m^2, R_{sm}^2, 0)$, the imaged circular points (ICP) that is its dual feature, are simply expressed as

$$\mathbf{I_{sm}} = \begin{bmatrix} R_m \\ iR_{sm} \\ 0 \end{bmatrix}, \mathbf{J_{sm}} = \begin{bmatrix} R_m \\ -iR_{sm} \\ 0 \end{bmatrix}.$$

Furthermore, we can assume that there is a physical camera to make the semi-metric image. This camera is referred to as a *semi-metric camera.* To find some properties of semi-metric cameras, image of absolute conic (IAC) of semi-metric cameras is studied.

Assuming that there are three vanishing points $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ whose directions are orthogonal in 3D, then the IAC would be expressed as [8]

$$\boldsymbol{\omega} = \alpha^2 \mathbf{l}_1 \mathbf{l}_1^\top + \beta^2 \mathbf{l}_2 \mathbf{l}_2^\top + \gamma^2 \mathbf{l}_3 \mathbf{l}_3^\top$$

where α, β and γ are proper scale factors and $\mathbf{l}_1$, $\mathbf{l}_2$, and $\mathbf{l}_3$ are vanishing lines given as

$$\mathbf{l}_1 = \mathbf{v}_1 \times \mathbf{v}_2, \mathbf{l}_2 = \mathbf{v}_2 \times \mathbf{v}_3, \mathbf{l}_3 = \mathbf{v}_3 \times \mathbf{v}_1.$$

In semi-metric space, the vanishing points $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ can be set as

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix},$$

which gives us the IAC in semi-metric space $\boldsymbol{\omega}_{sm}$ as

$$\boldsymbol{\omega}_{sm} = \begin{bmatrix} \beta^2 c^2 & 0 & -\beta ac \\ 0 & \gamma^2 c^2 & -\gamma bc \\ -\beta ac & -\gamma bc & \alpha^2 + \beta^2 a^2 + \gamma^2 b^2 \end{bmatrix}. \tag{1}$$

Because the ICPs are on the IAC,

$$\mathbf{I}_{sm}^\top \boldsymbol{\omega}_{sm} \mathbf{I}_{sm} = 0, \mathbf{J}_{sm}^\top \boldsymbol{\omega}_{sm} \mathbf{J}_{sm} = 0,$$

and we can find the relation that

$$\frac{{R_m}^2}{{R_{sm}}^2} = \frac{\gamma^2}{\beta^2}.$$

This means that the ratio of β and γ is equal to that of R_{sm} and R_m. By decomposing (1), the camera matrix in semi-metric space is given as

$$\mathsf{K}_{sm} = \begin{bmatrix} 1/R_{sm} & 0 & a \\ & 1/R_m & b \\ & & c \end{bmatrix} \tag{2}$$

up to scale.

As a consequence, the camera matrix K_{sm} represents a camera whose skew is zero, and its pixel aspect ratio is equal to a ratio between the aspect ratio of the reference rectangle R_m and the corresponding semi-metric aspect ratio R_{sm}. The principal point of the camera is expressed with the scaled third vanishing point $\mathbf{v}_3$ and the scale plays the role of a focal length. In other words, the semi-metric camera matrix is determined with scene information and a camera pose.

Naturally, the relation between IAC $\boldsymbol{\omega}$ in projective space and IAC $\boldsymbol{\omega}_{sm}$ in semi-metric space is obtained from basic conic transformation as

$$\mathsf{H}_{sm}^{-\top} \omega \mathsf{H}_{sm}^{-1} = \omega_{sm} \tag{3}$$

where H_{sm} is a plane homography from projective space to semi-metric space.

3 Estimation of Infinite Homography

In this section, we derive a parameterization of an infinite homography in terms of semi-metric warping matrices. Using the parameterization, it is possible to estimate an infinite homography linearly from images of two arbitrary rectangles.

3.1 Parameterization of Infinite Homography

Assuming that there are two views containing a projected unknown rectangle, then each semi-metric camera matrix K_{sm1}, and K_{sm2} would be expressed as

$$\mathsf{K}_{sm1} = \begin{bmatrix} 1/R_{sm} & 0 & s_1 m_1 \\ & 1/R_m & s_1 m_2 \\ & & s_1 m_3 \end{bmatrix}, \mathsf{K}_{sm2} = \begin{bmatrix} 1/R_{sm} & 0 & s_2 n_1 \\ & 1/R_m & s_2 n_2 \\ & & s_2 n_3 \end{bmatrix}$$

using (2). Note that we can set the value of R_{sm} to 1, as explained in Sect. 2. Furthermore, since the plane is identical, R_m is the same in both K_{sm1} and K_{sm2}.

A semi-metric image is generated by simple image warping. We can find the projection matrix of semi-metric camera directly as

$$\begin{aligned} \mathsf{P}_{sm} &= \mathsf{H}_{sm}\mathsf{K}\left[\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{r}_3\ \mathbf{t}\right] \\ &= \left[\mathsf{K}_{sm}\ \mathbf{e}_3\right] \\ &= \mathsf{K}_{sm}\left[\mathtt{I}_{3\times 3}\ \mathsf{K}_{sm}^{-1}\mathbf{e}_3\right] \end{aligned} \quad (4)$$

where $\mathbf{e}_3 = \left[0\ 0\ 1\right]^\top$. Note that all semi-metric cameras derived from an identical 3D rectangle are under pure translating motion. An infinite homography between two semi-metric cameras, K_{sm1} and K_{sm2} can be simply given as

$$\mathsf{T} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & t_z \end{bmatrix},$$

because an infinite homography is generally given as [1]

$$\mathsf{T} = \mathsf{K}_2\mathsf{R}_{21}\mathsf{K}_1^{-1} \quad (5)$$

and the two semi-metric cameras are under pure-translating, that means $\mathsf{R}_{21} = \mathtt{I}$. It gives us

$$\omega_{sm2} = \mathsf{T}^{-\top}\omega_{sm1}\mathsf{T}^{-1}$$

where ω_{sm1} and ω_{sm2} are IACs of the two semi-metric cameras.

Applying conic transformation given as (3) makes

$$\boldsymbol{\omega}_2 = \mathsf{H}_{sm2}^\top\mathsf{T}^{-\top}\mathsf{H}_{sm1}^{-\top}\boldsymbol{\omega}_1\mathsf{H}_{sm1}^{-1}\mathsf{T}^{-1}\mathsf{H}_{sm2},$$

and because the infinite homography H_∞^{12} from view 1 and view 2 transforms $\boldsymbol{\omega}_1$ to $\boldsymbol{\omega}_2$, the infinite homography is

$$\mathsf{H}_\infty^{12} = \mathsf{H}_{sm2}^{-1}\mathsf{T}\mathsf{H}_{sm1}. \quad (6)$$

This means that the infinite homography is expressed with semi-metric warping matrices H_{sm1} and H_{sm2} and the infinite homography T between two semi-metric cameras. Note that there are no camera assumptions such as static camera or zero-skew.

3.2 Linear Estimation of Infinite Homography

If a captured scene contains two arbitrary rectangles with an unknown aspect ratio, then the infinite homography is estimated linearly using the parameterization in (6).

Assuming that there are two views that contain two arbitrary rectangles named rectangle i and j, then we can find two infinite homographies with respect to two rectangles as

$$\mathsf{H}^{12}_{\infty,i} = \mathsf{H}^{-1}_{sm2,i}\mathsf{T}_i\mathsf{H}_{sm1,i}$$
$$\mathsf{H}^{12}_{\infty,j} = \mathsf{H}^{-1}_{sm2,j}\mathsf{T}_j\mathsf{H}_{sm1,j}$$

where $\mathsf{H}_{sm1,i}$ means a semi-metric warping matrix of view 1 w.r.t. the rectangle i.

However, the infinite homography is dependent only on the intrinsic parameters of the cameras and the relative rotation between two views. This means that the infinite homography is defined identically regardless of selecting which rectangle is used as a reference. This gives us a constraint equation of:

$$\rho\mathsf{H}^{-1}_{sm2,i}\mathsf{T}_i\mathsf{H}_{sm1,i} = \mathsf{H}^{-1}_{sm2,j}\mathsf{T}_j\mathsf{H}_{sm1,j} \tag{7}$$

where ρ is a proper scale factor.

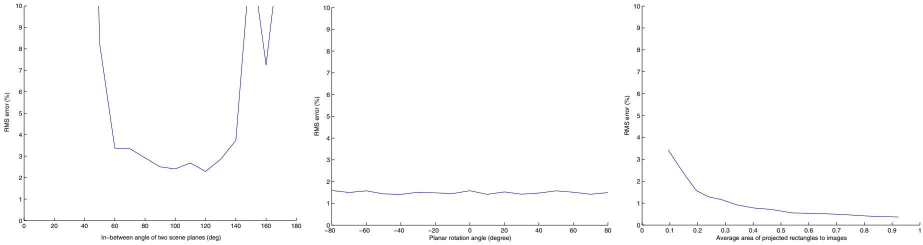
The unknowns are the parameters of T_i and T_j and a scale factor ρ. The number of unknowns is 7 and we have 9 equations, therefore we can easily solve the equation linearly. Note that we do not use any metric measurements, such as lengths or aspect ratios of the scene rectangles.

4 Application to Autocalibration

One of the most important applications of infinite homography is autocalibration of cameras [1]. If the infinite homographies between views captured by a static camera is known, then calibration can be possible linearly without any assumptions on cameras. We applied our proposed algorithm to the autocalibration of a static camera in order to provide validation.

The algorithm to build auto-calibration is as follows.

1. Track two arbitrary rectangles.
2. Find semi-metric warping matrices in all views w.r.t. the two rectangles.
3. Estimate proper transformation T_i and T_j between semi-metric space using (7).
4. Calculate the infinite homography H^{12}_{∞} with semi-metric transformation matrices and obtained proper transformation using (6).
5. Normalize the matrix so that $\det \mathsf{H}^{12}_{\infty} = 1$
6. Find the IAC using $\boldsymbol{\omega} = (\mathsf{H}^{12}_{\infty})^{-\top}\boldsymbol{\omega}(\mathsf{H}^{12}_{\infty})^{-1}$.
7. Determine the camera matrix K from the Cholesky decomposition $\boldsymbol{\omega} = (\mathsf{K}\mathsf{K}^{\top})^{-1}$.



(a) Pose differences **(b)** Planar rotation **(c)** Area

Fig. 3. Simulated performance of the proposed algorithm

This algorithm can be compared with previous works that uses information on scene geometry and proper camera assumptions [9, 8, 10, 11]. The key difference is that ours does not require any metric measurements from the scene, such as line lengths or aspect ratios of the rectangles. Furthermore, our algorithm does not contain camera assumptions, such as zero-skew or known aspect ratio of the pixels. Because it can be much easier to find some rectangles than to find some metrics in images, the proposed method is much more flexible than those given in the previous works.

We first analyzed the performance of the algorithm in various situations. We generated three views with two arbitrary rectangles in general poses and added Gaussian noises with a standard deviation of 0.5 to the corner of the rectangles. Fig. 3 depicts RMS errors of estimated focal length for 500 iterations. Fig. 3a shows the performance to pose differences between two planes in 3D. As expected, the algorithm become singular, when the in-between angle approaches to zero and 180 degrees, since it means the two rectangles are on an identical plane. In 40 degrees, one of the plane is orthogonal to the image plane, and all the features lie on a line. This is a singular case, and in other situation, the calibration is not much degraded for about 90 degrees. Fig. 3b shows the effects of the planar rotation of the world plane. We conclude that the direction of the model axis does not affect the performance of the algorithm. Fig. 3c shows the performance relative to the area of the rectangles used in the images. As expected, the performance of the algorithm increases with the rectangle size.

Fig. 4. Input images for auto-calibration using proposed method

The algorithm works well as long as the projected rectangles are larger than 10% of the whole images.

We next applied the algorithm to real images. Fig. 4 shows input images containing two arbitrary rectangles. The images were captured with a SONY DSC-F717 camera in 640×480 resolution. The exact values of the aspect ratios of the rectangles are unknown. Since the rectangles are placed arbitrarily, we cannot use the relative pose between two planes. Note that some imaged rectangles are rarely distorted projectively, so we cannot find the vanishing points or lines explicitly.

The estimated infinite homographies are

$$\mathsf{H}_\infty^{12} = \begin{bmatrix} 1.0406 & -0.0161 & -208.2218 \\ -0.0167 & 0.2692 & 864.6719 \\ 0.0004 & -0.0009 & 0.6885 \end{bmatrix},$$

$$\mathsf{H}_\infty^{12} = \begin{bmatrix} 1.0406 & -0.0161 & -208.2218 \\ -0.0167 & 0.2692 & 864.6719 \\ 0.0004 & -0.0009 & 0.6885 \end{bmatrix}$$

and

$$\mathsf{H}_\infty^{13} = \begin{bmatrix} 0.9991 & 0.1037 & -621.4391 \\ 0.0115 & 1.0388 & -127.6288 \\ 0.0006 & 0.0002 & 0.5807 \end{bmatrix}.$$

From the estimated infinite homographies, the intrinsic parameters of the camera is estimated as

$$\mathsf{K}_{estimated} = \begin{bmatrix} 899.4727 & 20.9762 & 322.9044 \\ 0 & 913.2549 & 297.9821 \\ 0 & 0 & 1.0000 \end{bmatrix}.$$

For comparison, we calibrated the camera using the well-known Zhang's plane based calibration method [12] with six metric planes as

$$\mathsf{K}_{Zhang} = \begin{bmatrix} 888.5763 & 14.3200 & 269.8877 \\ 0 & 887.2853 & 243.0086 \\ 0 & 0 & 1.0000 \end{bmatrix}.$$

Note that the proposed algorithm is a kind of autocalibration and does not require any kind of metric measurements, such as metric coordinates or line lengths. Furthermore, although we did not apply any robust method or refinement techniques such as non-linear minimization, the estimated camera parameters are comparable with only three images.

Figure 5 shows another real images captured with the SONY DSC-F717 camera. The yellow lines show the manually selected projected rectangles. The selected rectangles have different aspect ratios and the metric properties of each are unknown. We used only three images captured from different positions. Note

Fig. 5. Another input images for auto-calibration using proposed method

that there are little projective distortions on some projected rectangles. The estimated camera matrix is

$$\mathsf{K}_{estimated} = \begin{bmatrix} 728.5874 & 28.4393 & 352.8952 \\ 0 & 718.3721 & 285.1658 \\ 0 & 0 & 1.0000 \end{bmatrix}$$

and a result from Zhang's calibration method [12] with six metric planes is

$$\mathsf{K}_{Zhang} = \begin{bmatrix} 721.3052 & 2.7013 & 335.3498 \\ 0 & 724.9379 & 247.3248 \\ 0 & 0 & 1.0000 \end{bmatrix}.$$

This shows that autocalibration by our new proposed method can be applicable to real cameras by simply tracking two arbitrary rectangles in general poses.

5 Conclusion

In this paper, we have proposed a new method to estimate the infinite homography from images containing two arbitrary planar rectangles. The proposed method does not require any metric measurements, such as line lengths or aspect ratios of the rectangles.

To deal with rectangles efficiently, we introduce the concept of the semi-metric camera. Semi-metric cameras can be expressed with very simple forms of zero-skew cameras which are related with the aspect ratio of the scene rectangles and camera poses. Also the semi-metric cameras from general views that see an identical 3D rectangles can be regarded as pure-translating cameras. Using this formulation, an infinite homography between two views is expressed simply with semi-metric warping matrices and infinite homography between two semi-metric cameras. Using the fact that the infinite homographies derived from two different rectangles have to be identical, the unknown transformations are estimated linearly.

To validate our method, we used the proposed algorithm for autocalibration of a static camera. The autocalibration results obtained with our novel method were similar to those obtained with well-known plane-based calibration methods, even though our method required only four points on each rectangle and no further refinement.

Acknowledgments. This research has been partially supported by the Korean Ministry of Science and Technology for NRL Program(Grant number M1-0302-00-0064) and by Microsoft Research Asia.

References

1. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge (2000)
2. Criminisi, A., Reid, I., Zisserman, A.: Duality, rigidity and planar parallax. In Proceedings of European Conference on Computer Vision. (1998) II: 846
3. Rother, C., Carlsson, S.: Linear multi view reconstruction and camera recovery using a reference plane. International Journal of Computer Vision **49** (2002) 117–141
4. Irani, M., Anandan, P., Weinshall, D.: From reference frames to reference planes: Multi-view parallax geometry and applications. In Proceedings of European Conference on Computer Vision. (1998) II: 829
5. Irani, M., Anandan, P., Cohen, M.: Direct recovery of planar-parallax from multiple frames. IEEE Trans. Pattern Analysis and Machine Intelligence **24** (2002) 1528–1534
6. Pollefeys, M., Van Gool, L.: Stratified self-calibration with the modulus constraint. IEEE Trans. Pattern Analysis and Machine Intelligence **21** (1999) 707–724
7. Kim, J.-S., Kweon, I.S.: Semi-metric space: A new approach to treat orthogonality and parallelism. In Proceedings of Asian Conference on Computer Vision. (2005) To appear
8. Liebowitz, D., Zisserman, A.: Combining scene and auto-calibration constraints. In Proceedings of International Conference on Computer Vision. (1999) 293–300
9. Caprile, B., Torre, V.: Using vanishing points for camera calibration. International Journal of Computer Vision **4** (1990) 127–140
10. Liebowitz, D.: Camera Calibration and Reconstruction of Geometry from Images. PhD thesis, University of Oxford (2001)
11. Wilczkowiak, M., Boyer, E., Sturm, P.: Camera calibration and 3d reconstruction from single images using parallelepipeds. In Proceedings of International Conference on Computer Vision. (2001) I: 142–148
12. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Analysis and Machine Intelligence **22** (2000) 1330–1334