

Projective rectification from the fundamental matrix

John Mallon*, Paul F. Whelan

Vision Systems Group, Dublin City University, Dublin 9, Ireland

Received 29 January 2004; received in revised form 2 February 2005; accepted 30 March 2005

Abstract

This paper describes a direct, self-contained method for planar image rectification of stereo pairs. The method is based solely on an examination of the Fundamental matrix, where an improved method is given for the derivation of two projective transformations that horizontally align all the epipolar projections. A novel approach is proposed to uniquely optimise each transform in order to minimise perspective distortions. This ensures the rectified images resemble the original images as closely as possible. Detailed results show that the rectification precision exactly matches the estimation error of the Fundamental matrix. In tests the remaining perspective distortion offers on average less than one percent viewpoint distortion. Both these factors offer superior robustness and performance compared with existing techniques.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Projective rectification; Homography; Fundamental matrix; Distortion minimisation; Stereo vision

1. Introduction

Rectification is known to be a necessary step in stereoscopic analysis. The aligning of epipolar lines allows subsequent algorithms to take advantage of the epipolar constraint, reducing the search space to one dimension. In the uncalibrated case the resulting depth reconstruction is determined up to a projective transformation [1]. Many applications requiring such relative depth measures exist, including view synthesis [2] and robotic navigation [3].

It is known and easily demonstrated that the rectilinear mechanical alignment of two cameras is prohibitively difficult. The method developed in this paper is to simulate rectilinear images from those of arbitrarily placed cameras. This involves subjecting the images to a two dimensional projective transformation or planar homography. The homographies are calculated solely from an analysis of the Fundamental matrix, to re-orientate the epipolar projections parallel to the horizontal image axis. Undetermined parameters of the homographies operating on the x coordinate are then specified to maximise viewpoint

similarities between the original and rectified images, thus reducing distortional effects of the homographies and improving stereo matching. The rectification is therefore described by a reprojection onto two planes with relative degrees of freedom about the vertical axis only.

Projective rectification has many degrees of freedom. The problem is to find a rectification that minimises distortion of the rectified images. This work follows on from Hartley [4,1], where a rigid transformation is derived from the Fundamental matrix. This means that to first order, a points neighborhood undergoes rotation and translation only, hence the original and re-sampled images look similar. A related technique has been proposed by Al-Shalfan et al. [5], while Loop and Zhang [6] consider a stratified decomposition of the rectification homographies in order to minimise projective distortions. Pollefeys et al. [7] describe rectification as a reprojection onto a cylindrical surface instead of a plane, suitable for configurations when the epipole is within or close to an image. Papadimitriou and Dennis [8] present an approach for convergent stereo geometry, while Isgro and Trucco [9] consider rectification directly from point correspondences without explicitly determining the Fundamental matrix.

The main contribution of this paper is the proposal of a novel technique to reduce rectification distortions for the maximisation of viewpoint similarities between the original and rectified images. Previous distortion interpretations

* Corresponding author. Tel.: +353 1 7005869; fax: +353 1 7005508.
E-mail address: john.mallon@eeng.dcu.ie (J. Mallon).

have included orthogonality of image corners and maximising image content over the view window [3]. Loop and Zhang [6] consider distortion by attempting to force affine qualities on the homographies. As skew and aspect ratio are invariant to affine transforms, they make extra constraints upon the homographies to reduce distortion. Their approach is not optimal considering only one local region of the image. Our approach in contrast considers all regions of the image, enforcing first order orthogonal qualities in a natural way through Singular Value Decomposition.

The paper also presents an improved method for the computation of robust matching homographies, from a real Fundamental matrix estimated from noise affected points. This results in a rectification error equal to that of the Fundamental matrix error, significantly improving upon the alignment of epipolar lines compared to similar methods such as Hartley's [4,1] and Al-Shalfan et al. [5].

2. Background

We work only in two-dimensional projective space \mathcal{P}^2 , where points are represented by bold symbols e.g. \mathbf{p} . Coordinates of points are represented by 3 dimensional vectors, e.g. $\mathbf{p}=(u,v,w)^T$. If $w \neq 0$ then this represents the points in \mathcal{R}^2 expressed in Euclidean coordinates as $(u/w, v/w)^T$. When the scale has been fixed, i.e. $\mathbf{p}=(u/w, v/w, 1)^T$, these are known as affine points. If $w=0$, the points are known as points at ∞ , directions or affine vectors. Points are scale invariant in that $\mathbf{p}=\alpha\mathbf{p}$ ($\alpha \neq 0$). Lines are similarly represented by three-dimensional column vectors, e.g. $\mathbf{l}=(l_a, l_b, l_c)^T$. Transforms are 3×3 matrices of bold uppercase, e.g. \mathbf{T} , formed of columns $\mathbf{T}=[\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3]$ with entries $t_{11}, t_{12}, \dots, t_{33}$.

2.1. Epipolar geometry

Given two images of a scene, let \mathbf{m} and \mathbf{m}' be the projections of some 3D point \mathbf{M} in images \mathcal{I} and \mathcal{I}' , respectively. The intrinsic projective geometry between the two views is defined as:

$$\mathbf{m}'^T \mathbf{F} \mathbf{m} = 0, \quad (1)$$

where the Fundamental matrix \mathbf{F} [3,1,10] is a 3×3 matrix of rank 2. Given at least 8 point matches is possible to determine the matrix [1] [11]. The Fundamental matrix maps points in \mathcal{I} to lines in \mathcal{I}' upon which corresponding points lie. The image in \mathcal{I} of the camera center, C' , is termed epipole $\mathbf{e}=(e_u, e_v, 1)^T$ and similarly for \mathcal{I}' the image of C is $\mathbf{e}'=(e'_u, e'_v, 1)^T$:

$$\mathbf{F} \mathbf{e} = 0 = \mathbf{F}^T \mathbf{e}'.$$

The epipoles \mathbf{e} and \mathbf{e}' can be simply computed from the Singular Value Decomposition of $\mathbf{F}=\mathbf{U} \text{diag}(0, \sigma_1, \sigma_2) \mathbf{V}^T$ where $\mathbf{U}=[\mathbf{e}', \mathbf{u}_1, \mathbf{u}_2]$, $\mathbf{V}=[\mathbf{e}, \mathbf{v}_1, \mathbf{v}_2]$, σ_1 and σ_2 are the

typically non-zero singular values. All the epipolar lines in the respective images pass through the epipoles. In this paper we assume that the Fundamental matrix has been found, which requires at least 8 point matches for linear estimation. In addition we consider the origin of the images to be $\mathbf{c}=(0, 0)$, generally the top left corner.

3. Rectification

Image rectification is the process of re-aligning corresponding epipolar lines to become collinear and parallel with the x -axis as illustrated in Fig. 1. For a stereo sensor, mechanical adjustments of this calibre are difficult to achieve. However, given a description of the projective geometry between the two views, projective transformations can be applied resulting in rectified images. We choose the projective transformations uniquely to minimise distortions and maintain as accurately as possible the structure of the original images. This helps during subsequent stages, such as matching, ensuring local areas are not unnecessarily warped.

Rectification can be described by a transformation that sends the epipoles to infinity, hence the epipolar lines become parallel with each other. Additionally, we ensure that corresponding points have the same y coordinate by mapping the epipoles in the direction $\mathbf{e}=(1,0,0)^T$ or equivalently $\mathbf{e}=(e_u, 0, 0)^T$. The Fundamental matrix for such a rectified pair of images is:

$$\bar{\mathbf{F}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

The desired homographies give new image coordinates as $\bar{\mathbf{m}}=\mathbf{H}\mathbf{m}$ and $\bar{\mathbf{m}}'=\mathbf{H}'\mathbf{m}'$. It follows from Eq. (1) that $\bar{\mathbf{m}}'^T \bar{\mathbf{F}} \bar{\mathbf{m}} = 0$ and $\bar{\mathbf{m}}'^T \mathbf{H}'^T \mathbf{H}^T \bar{\mathbf{F}} \mathbf{H} \mathbf{m} = 0$ resulting in a set of constraints relating \mathbf{H} to \mathbf{H}' :

$$\mathbf{H}^T \bar{\mathbf{F}} \mathbf{H} = \mathbf{F}. \quad (2)$$

The homographies satisfying Eq. (2) are not unique, but similar to Hartley [4] we choose \mathbf{H} to transform the epipole \mathbf{e} to infinity:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ -e_v/e_u & 1 & 0 \\ -1/e_u & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ h_{21} & 1 & 0 \\ h_{31} & 0 & 1 \end{bmatrix}. \quad (3)$$

The determinant of the Jacobian, $\det(\mathbf{H})=1/(1-x/e_u)$, gives an indication of the changes or warping of local areas. At the origin the transformation appears orthogonal ($\det(\mathbf{H})=1$), while in general e_u is large in comparison to the image size. This ensures \mathbf{H} does not cause severe perspective distortion.

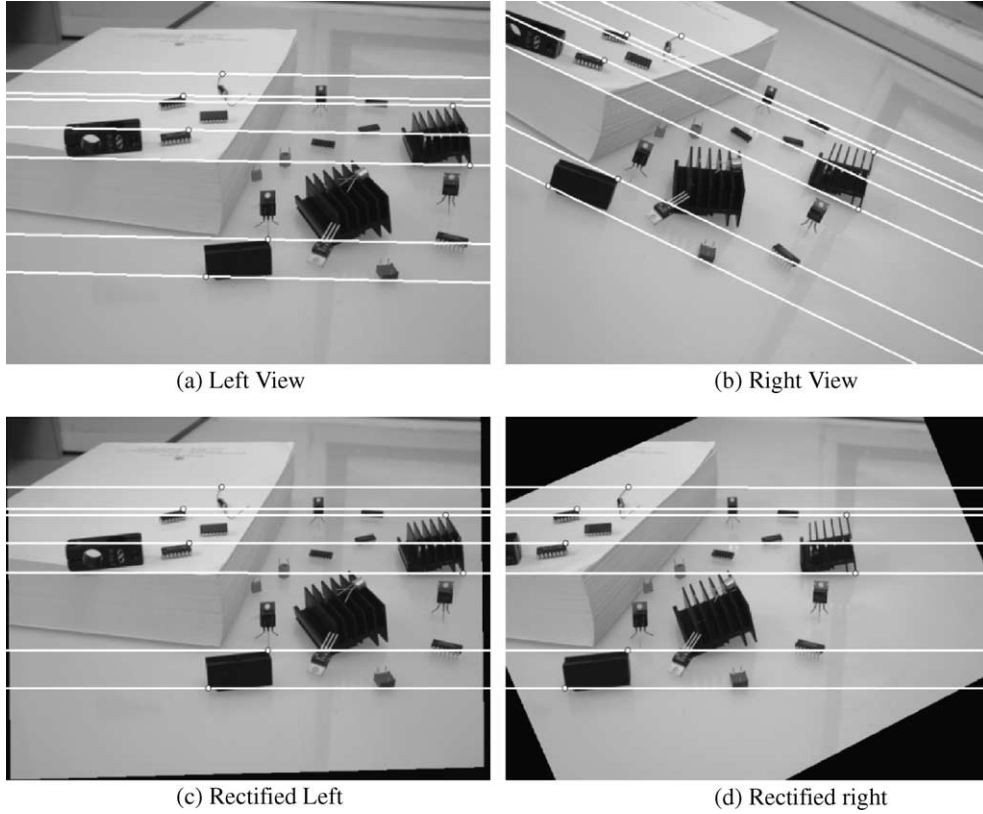


Fig. 1. Example of the rectification procedure. The original images are shown in (a) and (b) overlaid with their respective epipolar lines. After rectification these lines become collinear and parallel with the image x-axis, as shown (c) and (d).

3.1. Matching homography

Considering Eq. (2), it is clear that for an ideal Fundamental matrix there are no applicable constraints on the first row of \mathbf{H}' . Thus we specify the matching transformation \mathbf{H}' with the form:

$$\mathbf{H}' = \begin{bmatrix} 1 & 0 & 0 \\ h'_{21} & h'_{22} & h'_{23} \\ h'_{31} & h'_{32} & h'_{33} \end{bmatrix}.$$

Evaluating Eq. (2), we propose to estimate the entities of \mathbf{H}' by the elementary comparison of entries in Eq. (4), where α represents the (optional) arbitrary scale difference. The constraints on \mathbf{H}' are:

$$\begin{bmatrix} (h_{21}h'_{31} - h_{31}h'_{21}) & h'_{31} & -h'_{21} \\ (h_{21}h'_{32} - h_{31}h'_{22}) & h'_{32} & -h'_{22} \\ (h_{21}h'_{33} - h_{31}h'_{23}) & h'_{33} & -h'_{23} \end{bmatrix} = \alpha \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}. \quad (4)$$

Assuming an imperfect \mathbf{F} matrix, the solution for \mathbf{H}' can be robustly found in a least squares sense from Eq. (4) by the SVD of $\mathbf{B}\hat{\mathbf{p}} = 0$, where $\hat{\mathbf{p}} = (h'_{21}, h'_{22}, h'_{23}, h'_{31}, h'_{32}, h'_{33}, \alpha)^T$. Computing \mathbf{H}' in a least square sense using all the entries of \mathbf{F} significantly improves the rectification accuracy in

contrast with Hartley's [4] solution $\mathbf{H}' = \mathbf{H}([\mathbf{e}] \times \mathbf{F} + \mathbf{e}\mathbf{e}^T)$, and others (see Section 5).

4. Reducing rectification distortions

The application of \mathbf{H} and \mathbf{H}' does indeed rectify the images as required. However, as can be noted above, the first rows of the homographies are undetermined. This results naturally from the Fundamental matrix, which does not encapsulate any information about the position of the x coordinate. Weng et al. [12] shows that only one component of the image position of a point is used by the epipolar constraint. The projective transformations in general introduces distortions in the rectified images, specifically skewness and aspect/scale distortions. We are however free to specify the first rows of both homographies without invalidating the constraints we used to compute them, $\mathbf{A}\mathbf{H}\mathbf{e} = \mathbf{H}\mathbf{e}$, and similarly for the primed counterparts giving:

$$\mathbf{H}^T \mathbf{A}^T \bar{\mathbf{F}} \mathbf{A} \mathbf{H} = \mathbf{K}^T \bar{\mathbf{F}} \mathbf{K} = \mathbf{F}, \quad (5)$$

where $\mathbf{K} = \mathbf{A}\mathbf{H}$. \mathbf{A} and \mathbf{A}' are transformations of affine form:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The creation and loss of pixels as a result of the application of transformation \mathbf{K} , can be quantified in the local area of point \mathbf{p} by any norm of the Jacobian:

$$\mathbf{J}(\mathbf{K}, \mathbf{p}) = \begin{pmatrix} \frac{\partial \bar{x}}{\partial x} & \frac{\partial \bar{x}}{\partial y} \\ \frac{\partial \bar{y}}{\partial x} & \frac{\partial \bar{y}}{\partial y} \end{pmatrix}.$$

Let $\sigma_1(\mathbf{J})$ and $\sigma_2(\mathbf{J})$ be the non-zero singular values of \mathbf{J} in descending order. Ideally, an orthogonal transform that neither creates or destroys pixels will have singular values equal to one. In general $\sigma_1(\mathbf{J}) > 1$ for a transformation that overall creates extra pixels, and $\sigma_1(\mathbf{J}) < 1$ for an overall compression of pixels within a local region.

The search for the best compromise of the affine pair a_{11} and a_{12} to maintain orthogonality and perspective of the original image can thus be expressed by searching for the singular values that are closest to one. The Wielandt–Hoffman theorem [13] for singular values states that if \mathbf{A} and \mathbf{E} are matrices in $\mathbb{R}^{m \times n}$ with $m \geq n$, then:

$$\sum_{k=1}^n (\sigma_k(\mathbf{A} + \mathbf{E}) - \sigma_k(\mathbf{A}))^2 \leq \|\mathbf{E}\|_F^2,$$

where $\|\mathbf{E}\|_F$ is the Frobenius norm of \mathbf{E} . This indicates that if \mathbf{A} is perturbed by \mathbf{E} , the corresponding perturbation in any singular value of \mathbf{A} will be less than that of the Frobenius magnitude of \mathbf{E} . This means that the relationship between entries in a matrix and its singular values is a smooth function, making them very suitable for iterative search techniques.

4.1. Minimisation

The search is conducted by evaluating the singular values of the Jacobian at various points over the image. These points, \mathbf{p}_i can be simply specified as a grid covering the image area or as the corners of the image. The function to be minimised is then expressed as:

$$f(\hat{a}_{11}, \hat{a}_{12}) = \sum_{i=1}^n [(\sigma_1(\mathbf{J}(\mathbf{K}, \mathbf{p}_i)) - 1)^2 + (\sigma_2(\mathbf{J}(\mathbf{K}, \mathbf{p}_i)) - 1)^2]. \quad (6)$$

We minimise this functional using the Nelder and Mead simplex search method. Finite derivative methods can also be applied as the function inherently has smooth derivatives. Since a_{13} is an x direction shift it does not introduce any distortion. It can be chosen automatically to center the rectified image in the old image window if desired. The same procedures equally apply to the primed counterpart image. The rectification is determined solely on the estimate of the \mathbf{F} matrix. This has the advantage that no point correspondences are explicitly needed. Thus the rectification is invariant to the location or quantity of the point set, which overall tends towards a more consistent result.

5. Experiments

A selection of nine real examples are presented. The performance of the proposed rectification is quantified using various metrics, and compared side-by-side with two popular methods from the literature, Hartley's [4,1] and Loop and Zhang's [6]. The examples feature a range of Fundamental matrix accuracy levels, and requiring various transformation complexity. The images were taken with a digital camera with 640×480 pixel resolution, over random unknown baselines. The lens parameters remained unchanged throughout and lens distortion has been removed. The images are of indoor and outdoor scenes with relatively low and high depth of scene respectively. The Fundamental matrix was calculated using the linear normalised eight point method [1,11] using manually matched points. The examples are available at the VSG code archive web page, <http://www.eeng.dcu.ie/~vsl/vsgcode.html> including data.

5.1. Error metrics

The rectification technique is based solely on the estimation of the Fundamental matrix. Therefore, and according to Eq. (5), the rectification performance is directly related with the integrity of the Fundamental matrix. A direct evaluation of the accuracy of the Fundamental matrix is given by the perpendicular distance from a point to its epipolar line. Considering the corresponding points $\mathbf{m} = (u, v, 1)$ and $\mathbf{m}' = (u', v', 1)^T$, the epipolar line in \mathcal{G} is given by $\mathbf{l} = \mathbf{F}^T \mathbf{m}' = (l_a, l_b, l_c)^T$. The perpendicular line through \mathbf{m} is: $\mathbf{l}^\perp = (l_b, -l_a, (l_a v - l_b u))^T$ and the intersection point: $\mathbf{p}^\perp = \mathbf{l} \wedge \mathbf{l}^\perp = (u^\perp, v^\perp, 1)^T$, where \wedge is the cross-product. The Fundamental matrix error is the distance $E_f = ((u^\perp - u)^2 + (v^\perp - v)^2)^{1/2}$. The rectification precision is then evaluated as: $E_r = \|(\mathbf{K}\mathbf{m})_2^T - (\mathbf{K}'\mathbf{m}')_2^T\|$.

In general, it is not possible to avoid all distortions in a perspective transformation. Distortions in this case are defined as departures from the original image structure, such as skewness and relative scale changes. These factors can be quantified by measuring the proportional sizes and orthogonality of the transformed images. Thus $a = (w/2, 0, 1)^T$, $b = (w, h/2, 1)^T$, $c = (w/2, h, 1)^T$ and $d = (0, h/2, 1)^T$ are defined as four cross points of the image where w and h are the image width and height. These points are then transformed, whereupon the vectors $\bar{x} = \bar{b} - \bar{d}$ and $\bar{y} = \bar{c} - \bar{a}$ are formed. The orthogonality is then given as the angle of the upper left corner between the \bar{x} and \bar{y} vectors (ideally 90°). This angle is: $E_o = \cos^{-1}((\bar{x} \cdot \bar{y}) / (|\bar{x}| |\bar{y}|))$. The aspect ratio can be used to measure the relative size of the image. Taking the corner to corner length ratio by redefining the points $a = (0, 0, 1)^T$, $b = (w, 0, 1)^T$, $c = (w, h, 1)^T$ and $d = (0, h, 1)^T$, the vectors \bar{x} and \bar{y} are formed as before. The aspect ratio (ideally unity) is then defined as: $E_a = ((\bar{x}^T \bar{x}) / (\bar{y}^T \bar{y}))^{1/2}$.

5.2. Rectification performance

The rectification performance is concerned with quantifying only the y or row alignment of corresponding points over the images. Referring to Table 1, of interest is the Fundamental matrix error E_f , which is sufficiently characterised by its mean and standard deviation. As the rectification is only based on the Fundamental matrix, its error E_f represents the minimum expected rectification error E_r . The Fundamental error E_f ranges from 4.8(2.7) to 0.32(0.31) pixels in the examples. This is matched in every instance by the method proposed here, where E_r ranges from 4.5(2.4) to 0.11(0.32).

In comparison Hartley's [4,1] technique fails to match any of these results, with E_r ranging from 48.2(11.4) to 2.6(2.7) (see Table 1). It stems from a lack of robustness in the method used to match the homographies. In comparison to Loop and Zhang's methods [6], for some cases this rectification did not converge (nc), while convergence is questionable for the *Arch* example. As this technique is scale invariant, suitable scaling for each example was chosen manually. Disregarding the non-convergence cases the rectification error E_r ranges from 10.4(3.2) to 1.4(3.8). The poor alignment for these alternative techniques can

additionally be seen in the Figs. 2–5. It can be seen that the rectification process described here significantly outperforms the comparison techniques. Its convergence to the minimum error E_f in every case demonstrates good robustness.

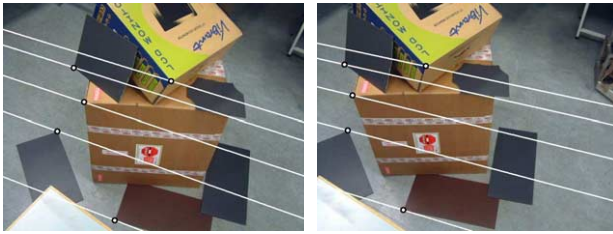
5.3. Distortion reduction

The rectification process introduces a necessary distortion to the images that realigns them horizontally relative to each other. We are free to specify the x position of pixels with an affine shearing transform that leaves the rectification unaffected. To this end we aim to preserve, as much as possible, the original viewpoint of each camera in order to avoid introducing unnecessary distortions. The distortion reduction criteria of Eq. (6) therefore strives to minimise local pixel warping throughout the image. Table 1 and Figs. 2–5 show the performance on the set of test images. The orthogonality E_0 and aspect ratio E_a are of interest and are calculated for both homographies. Orthogonality gives a intuitive measure of the distortion level. Taking the examples in Table 1, the average absolute orthogonal angle error for the proposed method is 0.8° for both \mathbf{H} and \mathbf{H}' .

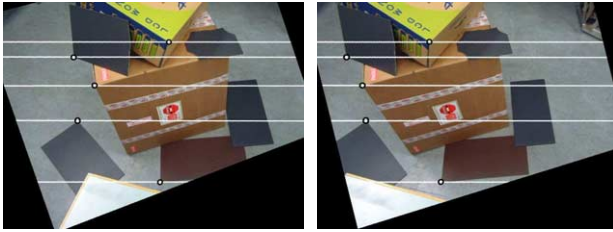
Table 1
Presents the error metrics described in Section 5.1 for nine different real world samples

Sample	F Mat. E_f mean (SD)	Method	Orthogonality E_0		Aspect ratio E_a		Rectification E_r	
			\mathbf{H}'	\mathbf{H}	\mathbf{H}'	\mathbf{H}	Mean	SD
Lab	4.861 (2.732)	Proposed	89.87	89.01	0.9960	0.9835	4.45	2.41
		Loop (nc)	–	–	–	–	–	–
		Hartley	99.07	96.56	1.1635	1.1111	23.02	4.53
Boxes*	0.5068 (0.3630)	Proposed	88.78	89.33	0.9785	0.9889	0.44	0.33
		Loop	97.77	95.69	1.1279	1.0900	4.35	9.20
		Hartley	86.56	94.99	0.9412	1.0846	33.36	8.65
Slates	0.5987 (0.4593)	Proposed	89.12	89.13	0.9852	0.9855	0.59	0.56
		Loop	37.29	37.15	0.2698	0.2805	1.14	3.84
		Hartley	89.96	88.54	1.0000	0.9769	2.27	5.18
Junk	0.3437 (0.2832)	Proposed	90.78	91.62	1.0233	1.0274	0.11	0.32
		Loop (nc)	–	–	–	–	–	–
		Hartley	102.67	99.60	1.3074	1.2466	14.38	8.51
Hall	1.9829 (1.2124)	Proposed	90.00	90.03	1.0003	1.0006	1.81	1.39
		Loop	91.14	91.58	1.0194	1.0271	4.92	2.40
		Hartley	102.56	90.48	1.2353	1.0081	2.59	2.75
Roof*	1.6422 (1.7085)	Proposed	88.35	88.23	1.1077	0.9700	1.96	2.95
		Loop	69.28	87.70	0.6665	1.0497	0.84	11.01
		Hartley	122.77	80.89	1.5256	0.8552	11.89	18.15
Arch	0.3244 (0.3123)	Proposed	91.22	90.26	1.0175	1.0045	0.22	0.33
		Loop	95.40	98.94	1.0991	1.1662	131.3	20.63
		Hartley	100.74	93.05	1.2077	1.0546	39.21	13.85
Yard*	0.6365 (0.4776)	Proposed	89.91	90.26	0.9987	1.0045	0.53	0.54
		Loop	133.62	134.27	2.1477	2.4045	8.91	13.19
		Hartley	101.95	91.91	1.2303	1.0335	48.19	11.49
Drive*	0.5684 (0.7568)	Proposed	90.44	90.12	1.0060	1.0021	0.18	0.91
		Loop	98.73	101.42	1.1541	1.2052	10.41	3.24
		Hartley	107.66	90.87	1.3491	1.015	3.57	3.43

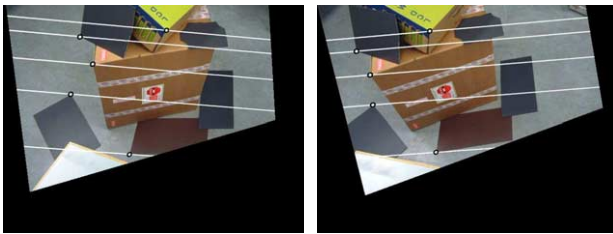
The errors are presented in mean (SD) format where applicable. Orthogonality is ideally 90° while the ideal aspect ratio is 1. The results are compared with two alternative techniques of Hartley [4,1] and Loop and Zhang [6], (nc)=no convergence. A selection of examples (Boxes, Roof, Yard and Drive) are shown in Figs. 2–5.



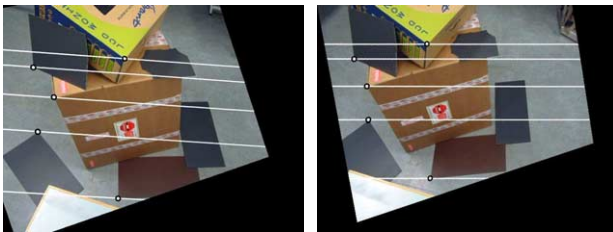
(a) Originals for Boxes example, Left and Right



(b) Proposed Rectification Method



(c) Loop and Zhang [6] Method



(d) Hartley [4] Method

Fig. 2. Boxes example including epipolar lines, see Table 1 for more details.

In comparison Hartley's [4,1] method introduces a significant quantity of distortion. This is expected as distortion is not considered, except to minimise disparity. As the results show this can introduce serious warping. The average absolute orthogonal angle error for \mathbf{H} and \mathbf{H}' from Table 1 is 4.2 and 12°. Loop and Zhang's methods [6] explicitly consider distortion, defined in a similar way to the proposed method, by preserving perpendicularity and aspect ratio. The results in Table 1 show that their method at no point matches our results. The average absolute orthogonal angle error for this method is 18 and 20° for \mathbf{H} and \mathbf{H}' . This is because their criteria is defined only for the midpoint of the image. An optimal estimate for one point does not mean it will be optimal for all image points, and indeed this is the case. In contrast our method considers the local areas over the entire image to preserve orthogonality giving superior results.



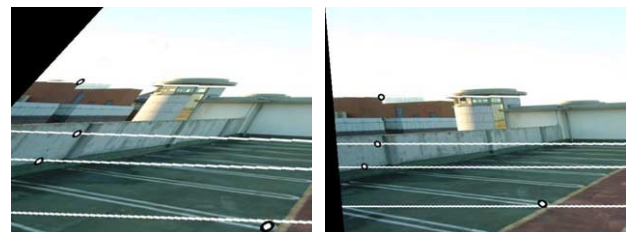
(a) Originals for Roof example, Left and Right



(b) Proposed Rectification Method



(c) Loop and Zhang [6] Method



(d) Hartley [4] Method

Fig. 3. Roof example including epipolar lines, see Table 1 for more details.

The rectification detailed above is based on the application of planar 2D projective transformations. As evident from Section 3, these techniques are not applicable for configurations where the epipole is within an image. In this case it is possible to cause severe image distortions, even splitting connected regions by their application. In general, the geometry that results in such undesirable situations, such as forward translation, is not generally encountered in a stereo-like setup.

6. Conclusion

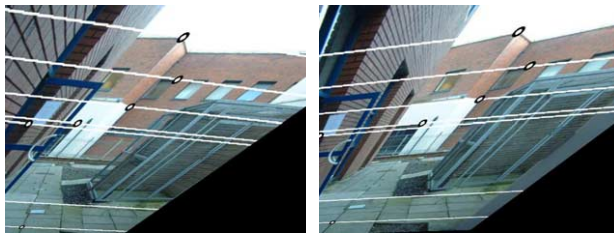
This paper describes a robust method for uncalibrated planar rectification for a pair of stereo images taken from distinct viewpoints. It is simple to implement and based



(a) Originals for Yard example, Left and Right



(b) Proposed Rectification Method



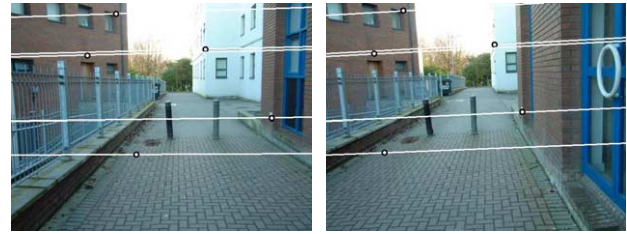
(c) Loop and Zhang [6] Method



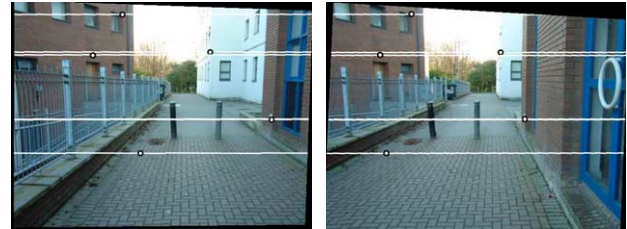
(d) Hartley [4] Method

Fig. 4. Yard example including epipolar lines, see Table 1 for more details.

solely on the estimated Fundamental matrix. A much improved method is given for the computation of matching perspective transformations, with experimental results showing that the rectification accuracy, or epipolar alignment, is equal to the error in the Fundamental matrix estimation. A novel technique is described to reduce the inevitable perspective distortions. This ensures that the rectified images resemble the originals as closely as possible, virtually eliminating unnatural skews and scaling. This has clear advantages for subsequent processing steps. The distortion minimisation is carried out by searching through a SVD for the best first order approximation of an orthogonal-like transform throughout the image window. Detailed comparison results clearly indicate much improved performance for both the rectification process and the distortion reduction techniques.



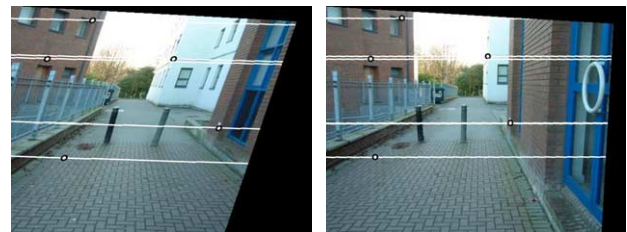
(a) Originals for Drive example, Left and Right



(b) Proposed Rectification Method



(c) Loop and Zhang [6] Method



(d) Hartley [4] Method

Fig. 5. Drive example including epipolar lines, see Table 1 for more details.

Acknowledgements

We gratefully acknowledge Prof. John Barron (University of Western Ontario) for his comments and discussions, and the referees for their comments.

References

- [1] R.I. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, second ed., Cambridge University Press, 2003.
- [2] K.C. Ng, M. Trivedi, H. Ishiguro, Generalized multiple baseline stereo and direct virtual view synthesis using range-space search, match, and render, *International Journal of Computer Vision* 47 (1–3) (2002) 131–147.
- [3] O. Faugeras, Q.-T. Luong, *The Geometry of Multiple Images*, MIT Press, Cambridge, MA, 2001.
- [4] R.I. Hartley, Theory and practice of projective rectification, *International Journal of Computer Vision* 35 (2) (1999) 115–127.

- [5] K.A. Al-Shalfan, J.G.B. Haigh, S.S. Ipson, Direct algorithm for rectifying pairs of uncalibrated images, *Electronics Letters* 36 (5) (2000) 419–420.
- [6] C. Loop, Z. Zhang, Computing rectifying homographies for stereo vision, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 1999, pp. 125–131.
- [7] M. Pollefeys, R. Koch, L.V. Gool, A simple and efficient rectification method for general motion, in: *Proceedings of the International Conference on Computer Vision*, vol. 1, 1999, pp. 496–501.
- [8] D.V. Papadimitriou, T.J. Dennis, Epipolar line estimation and rectification for stereo image pairs, *IEEE Transactions on Image Processing* 5 (4) (1996) 672–676.
- [9] F. Isgro, E. Trucco, On projective rectification, in: *Proceedings IEE Conference on Image Processing and Analysis*, 1999, pp. 42–46.
- [10] X. Armangué, J. Salvi, Overall view regarding fundamental matrix estimation, *Image and Vision Computing* 21 (2003) 205–220.
- [11] R.I. Hartley, In defense of the 8-point algorithm, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (6) (1997) 580–593.
- [12] J. Weng, N. Ahuja, T.S. Huang, Optimal motion and structure estimation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15 (9) (1993) 864–884.
- [13] G.H. Golub, C.F.V. Loan, *Matrix Computation*, third ed., Johns Hopkins University Press, 1996.