Uncalibrated three-view image rectification

Conference Paper *in* Image and Vision Computing · December 2000

Impact Factor: 1.59

1 author:

Changming Sun

The Commonwealth Scientific and Industrial …

125 PUBLICATIONS **1,549** CITATIONS

SEE PROFILE

ELSEVIER

Image and Vision Computing 21 (2003) 259–269

image AND vision COMPUTING

www.elsevier.com/locate/imavis

Uncalibrated three-view image rectification

Changming Sun

CSIRO Mathematical and Information Sciences, Locked Bag 17, North Ryde, NSW 1670, Australia

Received 23 October 2002; accepted 24 October 2002

Abstract

Image rectification is a process of transforming a set of images into a new set such that the epipolar lines in the transformed images have the same direction as the image rows or columns to enable an efficient and reliable stereo matching. Previous algorithms for stereo image rectification either work for two view uncalibrated or two/three view calibrated situations. In this paper we propose several novel techniques to rectify uncalibrated trinocular images using the trilinear tensor or projective invariants or fundamental matrices obtained from a triplet of images. Our new methods include: a rotation and skew method, an affine transformation method, and a vanishing points method. Real images have been used for testing purposes, and accurate results have been obtained.

Keywords: Rectification; Uncalibrated images; Trilinear tensor; Fundamental matrix; Epipolar lines; Epipoles; Vanishing points; Vanishing line; Projection matrix

1. Introduction

Image rectification is an important step in the three dimensional analysis of scenes. For stereo vision, image rectification can increase both the reliability and the speed of the disparity estimation process. This is because in the rectified images, the relative rotation among the original images have been removed and the disparity search happens along the image horizontal or vertical scanlines. The rectification process requires certain camera calibration parameters or weakly calibrated (uncalibrated) epipolar geometries of the image pair or image triplet.

Hartley gave a mathematical basis and a practical algorithm for the rectification of stereo images from widely different viewpoints [1]. Al-Shalfan et al. presented a direct algorithm for rectifying pairs of uncalibrated images [2]. Isgrò and Trucco presented a robust algorithm performing uncalibrated rectification which does not require explicit computation of the epipolar geometry [3]. Pollefeys et al. proposed a simple and efficient rectification method for general two view stereo images [4]. Loop and Zhang proposed a technique for computing rectification homographies for stereo vision [5]. Papadimitriou and Dennis presented an algorithm for rectifying stereo images when the images are taken with convergence geometry (coplanar X and Z axes and parallel Y axes) [6]. The rectification processes by Hartley, Al-Shalfan et al., Isgrò and Trucco, Pollefeys et al., Papadimitriou and Dennis and Loop and Zhang can work for uncalibrated cameras. These algorithms are developed for two view rectification situations.

Ayache and Hansen presented a technique for calibrating and rectifying pairs or triplets images [7]. In their case, a camera matrix needs to be estimated. Therefore the algorithm works for calibrated cameras. Shao and Fraser also developed a rectification method for calibrated trinocular cameras [8]. The Digiclops developed at Point Grey Research Inc used three calibrated cameras for stereo vision after rectification [9]. Fusiello et al., presented a compact algorithm for rectifying calibrated stereo images [10]. All these algorithms only work for calibrated cameras.

In this paper, we propose several novel algorithms for rectifying three uncalibrated images. Section 2 describes our new methods for rectifying the trinocular images. Section 3 shows the experimental results obtained using our new rectification methods. Section 4 gives concluding remarks.

2. Rectifying uncalibrated trinocular images

2.1. Obtaining fundamental matrices

In the case of three cameras or three images taken by one camera, it has been shown in [11] that the correspondence constraint is expressed by the trilinear tensor. The tensor can

E-mail address: changming.sun@csiro.au.

0262-8856/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0262-8856(02)00157-9

be recovered linearly from at least seven corresponding points or lines across the three views. It is also shown in [11] that the concatenation of epipolar geometries across three views fails in cases where trilinearities do not. The trilinearities use all the three views together, rather than in pairs as in the case of the fundamental matrix, thereby gaining additional numerical stability.

Introduce a Cartesian coordinate system with the origin at the pin-hole of the camera and the z-axis aligned with the optical axis and pointing outward. If we denote a matching point among the three views as $\mathbf{p}_i^{(1)} = (x_i^{(1)}, y_i^{(1)}, 1)^T$, $\mathbf{p}_i^{(2)} = (x_i^{(2)}, y_i^{(2)}, 1)^T$, and $\mathbf{p}_i^{(3)} = (x_i^{(3)}, y_i^{(3)}, 1)^T$, we have a matching set $\{(\mathbf{p}_i^{(1)}, \mathbf{p}_i^{(2)}, \mathbf{p}_i^{(3)}), i = 0, \cdots, N-1\}$ where N is the number of matched points appearing in all three views. If we denote the trilinear tensor as $\mathbf{T}'_{\alpha\beta\gamma}$, ($\mathbf{T}'$ is a $3 \times 3 \times 3$ tensor, $\alpha, \beta, \gamma = 1, 2, 3$), the relationships which governs the parameters of the trilinear tensor and a matching point can be written as the following linear equations [11]:

$$x_i^{(3)}\mathbf{T}'_{\alpha 13}\mathbf{p}_i^{(1)} - x_i^{(3)}x_i^{(2)}\mathbf{T}'_{\alpha 33}\mathbf{p}_i^{(1)} + x_i^{(2)}\mathbf{T}'_{\alpha 31}\mathbf{p}_i^{(1)} - \mathbf{T}'_{\alpha 11}\mathbf{p}_i^{(1)} = 0,$$

$$y_i^{(3)}\mathbf{T}'_{\alpha 13}\mathbf{p}_i^{(1)} - y_i^{(3)}x_i^{(2)}\mathbf{T}'_{\alpha 33}\mathbf{p}_i^{(1)} + x_i^{(2)}\mathbf{T}'_{\alpha 32}\mathbf{p}_i^{(1)} - \mathbf{T}'_{\alpha 12}\mathbf{p}_i^{(1)} = 0,$$

$$x_i^{(3)}\mathbf{T}'_{\alpha 23}\mathbf{p}_i^{(1)} - x_i^{(3)}y_i^{(2)}\mathbf{T}'_{\alpha 33}\mathbf{p}_i^{(1)} + x_i^{(2)}\mathbf{T}'_{\alpha 31}\mathbf{p}_i^{(1)} - \mathbf{T}'_{\alpha 21}\mathbf{p}_i^{(1)} = 0,$$

$$y_i^{(3)}\mathbf{T}'_{\alpha 23}\mathbf{p}_i^{(1)} - y_i^{(3)}y_i^{(2)}\mathbf{T}'_{\alpha 33}\mathbf{p}_i^{(1)} + x_i^{(2)}\mathbf{T}'_{\alpha 32}\mathbf{p}_i^{(1)} - \mathbf{T}'_{\alpha 22}\mathbf{p}_i^{(1)} = 0$$

where $\mathbf{T}'_{\alpha **}$ is a three element vector with $\alpha = 1,2,3$. i is the index for the i^{th} corresponding point among all the three views.

There are twenty-seven (27) parameters to be estimated for the trilinear tensor governing the relationships between each of the triplet images. For each matching point, four equations can be obtained. If at least seven matched points are available, twenty-eight equations can be established with 27 unknown parameters for the trilinear tensor. The tensor can therefore be estimated by solving the set of linear equations. Normalisation of image points can be applied to improve the condition of the problem [12].

It has been shown that the fundamental matrices which govern the epipolar geometry between two views can be obtained from the trilinear tensor. Shashua and Werman gave an equation that shows the relationship between the fundamental matrix $\mathbf{F}_{12}$ (the fundamental matrix between view 1 and view 2) and the obtained trilinear tensor $\mathbf{T}'$ as [13]:

$$\mathbf{F}_{12}^T\mathbf{E}_\beta + \mathbf{E}_\beta^T\mathbf{F}_{12} = 0, \qquad \beta = 1, 2, 3 \tag{1}$$

where the matrix $\mathbf{E}_\beta = \mathbf{T}'_{*\beta *}$. Each $\mathbf{E}_\beta$ is a 3×3 matrix. Since the left hand side of the Eq. (1) is a symmetric matrix, for each β, we have 6 independent equations relating $\mathbf{F}_{12}$ and the obtained trilinear tensor $\mathbf{T}'$. $\mathbf{F}_{12}$ can therefore be obtained using least squares method because we have 18 homogeneous equations for 9 unknowns. The fundamental matrix $\mathbf{F}_{12}$ thus obtained governs the relationship between view 1 and view 2. Similarly we can have the following relationship:

$$\mathbf{F}_{13}^T\mathbf{E}_\alpha + \mathbf{E}_\alpha^T\mathbf{F}_{13} = 0, \qquad \alpha = 1, 2, 3 \tag{2}$$

where matrix $\mathbf{E}_\alpha = \mathbf{T}'_{\alpha **}$. From the above equation, we can obtain the fundamental matrix $\mathbf{F}_{13}$ which is related to view 1 and view 3. Therefore we have obtained two fundamental matrices that govern the epipolar geometries for views 1 and 2 and views 1 and 3.

The method described above for obtaining the fundamental matrices via the trilinear tensor involves only solving a set of linear equations. The trilinear tensor is also obtained by solving a set of linear equations. Quan developed algorithms for the computation of the invariants of sets of six points from three images [14]. Quan's method only requires six matched points appearing in three images while the trilinear tensor method needs seven. In Quan's method, the invariants are calculated first, and then camera projection matrices are determined. The epipolar geometry can be obtain by using the obtained camera projection matrices.

The fundamental matrices can also be estimated directly from matching points between two views [15]. Once the fundamental matrices among the three views are available, they can then be used for rectification purposes.

In practice, the matching points among the views obtained by automatic matching algorithms may contain mismatches. Therefore robust estimation methods such as those in [16] can be used. Nonlinear constraints can also be used for estimating the trilinear tensor [17].

2.2. Projection matrix for the reference view

Given the epipolar geometry defined by the fundamental matrix $\mathbf{F}_{12}$, a pair of epipoles can be obtained by $\mathbf{F}_{12}\mathbf{e}_{12} = 0$ and $\mathbf{e}_{21}^T\mathbf{F}_{12} = 0$, where $\mathbf{e}_{12}$ and $\mathbf{e}_{21}$ are the two epipoles on images 1 and 2 based on the fundamental matrix $\mathbf{F}_{12}$. Similarly, two epipoles ($\mathbf{e}_{13}$ and $\mathbf{e}_{31}$) on images 1 and 3 based on the fundamental matrix $\mathbf{F}_{13}$ can also be obtained. The image rectification process is to transform the images such that the epipoles in the transformed images are at the infinity either on the x-axis or on the y-axis. In the usual two view stereo case, the epipoles of the transformed images are at the infinity point $(1,0,0)^T$. For three view stereo, the epipoles for the transformed left and right image pair will be at $(1,0,0)^T$. However, the epipoles for the transformed left and top image pair will be at $(0,1,0)^T$ (on the y-axis). Here the terms 'left', 'right' and 'top' refer to the relative positions only. The left image is also the reference image. Fig. 1 illustrates the initial positions of the epipoles obtained in the trinocular images. Note that the positions of these epipoles are mostly outside the input images although we draw them inside the image.

We need to find a mapping function which will transform the two epipoles $\mathbf{e}_{12}$ and $\mathbf{e}_{13}$ on the original reference image into two points at the infinity $(1,0,0)^T$ and

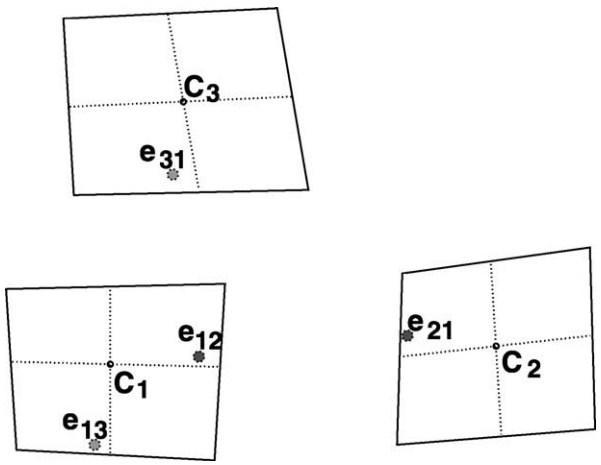


Fig. 1. The initial positions of the epipoles in the three images. C_1 indicates the image centre for the left or the reference image. C_2 indicates the image centre for the right image. C_3 indicates the image centre for the top image. e_{12} and e_{13} are the epipoles in the reference image. e_{21} is the epipole in the right image and e_{31} is the epipole in the top image.

$(0,1,0)^T$ in the rectified image. In order to reduce image distortions, image transformations need to be as rigid as possible. In the following subsections, we will present several novel algorithms for rectifying the reference image.

2.2.1. Rotation and skew method

Assuming that the image centre is at $(u,v)^T$, we can use the following transformation to shift the image coordinate system to the image centre:

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & -u \\ 0 & 1 & -v \\ 0 & 0 & 1 \end{pmatrix}.$$

Then the image can be rotated such that the epipole $\mathbf{e}_{12} = (e_{12}[0], e_{12}[1], 1)^T$ lies on the x-axis. This rotation transformation takes the form of:

$$\mathbf{R} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where $\theta = \arctan(e_{12}[1]/e_{12}[0])$. After the epipole $\mathbf{e}_{12}$ has been transformed to lie on the x-axis, we need to move epipole $\mathbf{e}_{13}$ on the y-axis. This can be obtained by a skew operation:

$$\mathbf{S} = \begin{pmatrix} 1 & s & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

where $s = -t[0]/t[1]$, and $\mathbf{t} = (t[0], t[1], 1)^T$ is the transformed position of $\mathbf{e}_{13}$ after applying the $\mathbf{T}$ and $\mathbf{R}$ transformations. Now the two epipoles have been transformed on to the image axes with value k_x on the x-axis for the first epipole and value k_y on the y-axis for the second epipole. The next step will be to shift the epipole positions to the infinity. This transformation can be achieved using the following matrix:

$$\mathbf{K} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/k_x & -1/k_y & 1 \end{pmatrix}.$$

The combined transformation matrix is:

$$\mathbf{P}_1 = \mathbf{KSRT}. \tag{4}$$

This transformation will have the effect of shifting the two epipoles in the reference image to the infinity on two orthogonal coordinate axes. Note that the values of θ, s, k_x and k_y are obtained from the immediate previous transformation operation. This sequence of transformation to move the epipoles to the infinity is shown in Fig. 2.

We can also apply the skew transformation immediately after the $\mathbf{T}$ transformation so that the two vectors formed from the image centre and the two epipoles are perpendicular to each other. The parameter s for the skew transformation in this case can be obtained using the method to be described in Section 2.2.3. Then a rotation transformation can be applied to make the epipoles lie on the image axes. This is then followed by the $\mathbf{K}$ transformation. The combined transformation process will then be:

$$\mathbf{P}_1 = \mathbf{KRST}. \tag{5}$$

2.2.2. Affine transformation method

An alternative way of aligning the epipoles in the reference image to the image axes is to choose firstly an affine transformation such that the two epipoles are on

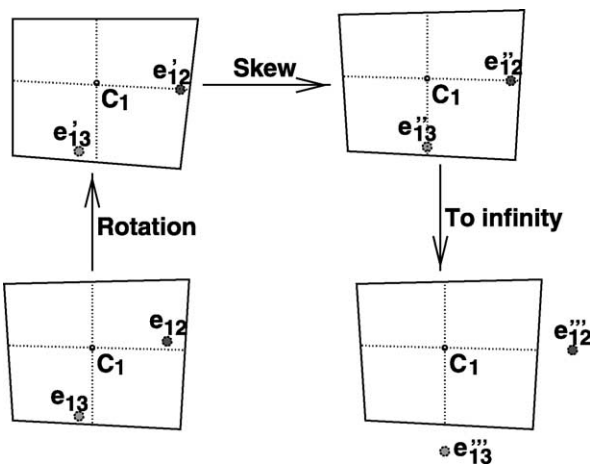


Fig. 2. The transformation process for moving the epipoles in the reference image to the infinity. The rotation transformation moves epipole e_{12} on to the x-axis. The skew transformation moves the epipole e'_{13} on to the y-axis. Then the $\mathbf{K}$ transformation moves the two epipoles to the infinity.

the images axes. The general affine transformation is given by:

$$\begin{cases} x' = a_{11}x + a_{12}y + a_{13} \\ y' = a_{21}x + a_{22}y + a_{23} \end{cases}$$

One needs to have three pair of point positions before and after the transformation to determine the affine parameters. We can use the image centre and the two epipoles for this purpose. If we assume that the image centre does not change for this affine transformation, the parameters a_{13} and a_{23} need to be zero. Then we can use the two epipole positions to determine a_{11}, a_{12}, a_{21} and a_{22} using the following transformation:

$$\begin{cases} x' = a_{11}x + a_{12}y \\ y' = a_{21}x + a_{22}y \end{cases}$$

The position for the first epipole $\mathbf{e}_{12} = (e_{12}[0], e_{12}[1], 1)^T$ will be transformed to $\mathbf{e}'_{12} = (e'_{12}[0], 0, 1)^T$, where $e'_{12}[0] = \|\mathbf{e}_{12}\|$. Similarly, the position for the second epipole $\mathbf{e}_{13} = (e_{13}[0], e_{13}[1], 1)^T$ will be transformed to $\mathbf{e}'_{13} = (0, e'_{13}[1], 1)^T$, where $e'_{13}[1] = \|\mathbf{e}_{13}\|$. By solving two sets of linear equations, the affine parameters can be obtained as:

$$\begin{cases} a_{11} = -\|\mathbf{e}_{12}\| \cdot e_{13}[1]/det \\ a_{12} = +\|\mathbf{e}_{12}\| \cdot e_{13}[0]/det \\ a_{21} = +\|\mathbf{e}_{13}\| \cdot e_{12}[1]/det \\ a_{22} = -\|\mathbf{e}_{13}\| \cdot e_{12}[0]/det \end{cases}$$

where $det = e_{13}[0] \cdot e_{12}[1] - e_{12}[0] \cdot e_{13}[1]$. After obtaining the affine parameters, we have the following affine transformation matrix $\mathbf{A}$ to move the epipoles on to the image axes:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

After these, the $\mathbf{K}$ transformation can be used to bring the epipoles to the infinity. By using the affine transformation method, we have the combined transformation matrix:

$$\mathbf{P}_1 = \mathbf{KAT}. \tag{6}$$

Fig. 3 illustrates the process of moving the epipoles to the infinity using the affine transformation methods.

2.2.3. Vanishing points method

Vanishing points have been used for 3D reconstruction from images, matching perspective views [18], shape from texture [19], and plane rectification [20,21].

Here we draw the connection that epipoles in the original image space are vanishing points for parallel lines in the rectified image space. We use the vanishing points property in this section for a new application – trinocular image rectification. In our situation, we have two epipoles

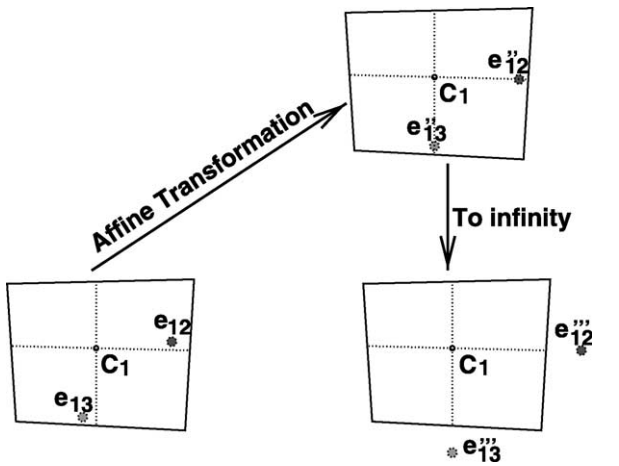


Fig. 3. Another transformation process to move the epipoles to the infinity. The affine transformation moves the two epipoles on to the two image axes. Then the $\mathbf{K}$ transformation move the epipoles to the infinity.

on the image plane of the reference image. These two epipoles, which are treated as vanishing points (as epipolar lines passing through epipoles), define the vanishing line $\mathbf{l} = (l_1, l_2, l_3)^T$. The following perspective projection matrix which is constructed from the vanishing line $\mathbf{l}$ transforms parallel lines into converging lines [19]:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{pmatrix}. \tag{7}$$

After applying the projective transformation $\mathbf{P}$ on the original image, the epipolar lines passing through each of the epipole become parallel, and the positions of the two epipoles are at the infinity. The positions of the transformed epipoles after the $\mathbf{T}$ and $\mathbf{P}$ transformation are at $\mathbf{m} = \mathbf{PTe}_{12} = (u_1, v_1, 0)^T$ and $\mathbf{n} = \mathbf{PTe}_{13} = (u_2, v_2, 0)^T$. The two sets of epipolar lines are usually not perpendicular to each other after this perspective transformation. We then need to apply a skew transformation such that the two sets of epipolar lines are orthogonal to each other.

The parameter s in the skew matrix needs to be estimated using the condition that the skew transformation makes the two vectors $\mathbf{m}$ and $\mathbf{n}$ orthogonal to each other. The condition says that the dot product of $(\mathbf{Sm})$ and $(\mathbf{Sn})$ needs to be zero. That is:

$$(\mathbf{Sm}) \cdot (\mathbf{Sn}) = 0 \tag{8}$$

where $\mathbf{S}$ is defined in Eq. (3). Expanding Eq. (8) we obtain

$$v_1 v_2 s^2 + (u_1 v_2 + u_2 v_1)s + u_1 u_2 + v_1 v_2 = 0.$$

Using the known positions of $\mathbf{m}$ and $\mathbf{n}$, we can obtain the solutions for s from the above equation. There are usually two solutions for s. The one with the smaller absolute value should be used in order to reduce the amount of image transformation.

After applying the $\mathbf{SPT}$ transformation, each set of epipolar lines become parallel, and the two sets of epipolar

lines are orthogonal to each other. We then need to rotate the image so that the epipolar lines are along the image axes. This rotation matrix **R** is obtained using a similar procedure as described in Section 2.2.1. Therefore the combined transformation for the reference image can be obtained as

$$\mathbf{P}_1 = \mathbf{RSPT}. \tag{9}$$

2.2.4. Optimising image translation

The u and v values mentioned at the beginning of Section 2.2.1 are taken as the centre of the reference image, i.e. u equals half of the image column numbers and v equals half of the image row numbers. The rectified reference image obtained by applying the transformation matrix $\mathbf{P}_1$ is usually different from the original image. Our intention is to minimise the image difference before and after the rectification process.

We can apply a minimisation process to search for the u and v values so that the image difference before and after the rectification process is minimised. The image difference is calculated as the sum of the distances for the four corners of the rectified and the original images. For each iteration of the minimisation process, a pair of intermediate u and v values is used for obtaining the projection matrix $\mathbf{P}_1$ which is used for transforming the four corners of the original image. Then the sum of the distances for the four corners of the rectified and the original images can be calculated.

After the minimisation process, a pair of u and v values (which defines **T**) and also the $\mathbf{P}_1$ matrix are obtained. This $\mathbf{P}_1$ matrix will be used to rectify the reference image.

We can carry out another search of u and v after the $\mathbf{P}_1$ matrix has been fixed. For this round of optimisation, the $\mathbf{P}_1$ remain unchanged, but just search through the u and v values so that the sum of the distances for the four corners of the rectified and the original images is minimised. Fig. 4 shows the effects of adjusting the u,v values to minimise the changes of the rectified images. Fig. 4(a) is the result obtained without the u,v optimisation. Fig. 4(b) shows the result obtained with the u,v optimisation. There are more black pixels in Fig. 4(a) than in Fig. 4(b). The original input images are shown in Fig. 6.

2.3. Projection matrices for two other views

Given the projection matrix $\mathbf{P}_1$ for the reference view and the two fundamental matrices $\mathbf{F}_{12}$ and $\mathbf{F}_{13}$, the transformation for the other two views can be obtained. Let $\mathbf{P}_2$ and $\mathbf{P}_3$ be the transformation matrices to rectify view 2 and view 3 respectively. It is proved in [1] that if $\mathbf{F}_{12}$ is factorised as $[\mathbf{e}_{12} \times \mathbf{M}_{12}]$, where $\mathbf{M}_{12}$ is a non-singular matrix, then

$$\begin{aligned}\mathbf{P}_2 &= (\mathbf{I} + \mathbf{P}_1\mathbf{e}_{12}\mathbf{a}^T)\mathbf{P}_1\mathbf{M}_{12} \\ &= (\mathbf{I} + (1\ 0\ 0)^T\mathbf{a}^T)\mathbf{P}_1\mathbf{M}_{12} = \mathbf{C}_2\mathbf{P}_1\mathbf{M}_{12}\end{aligned}$$

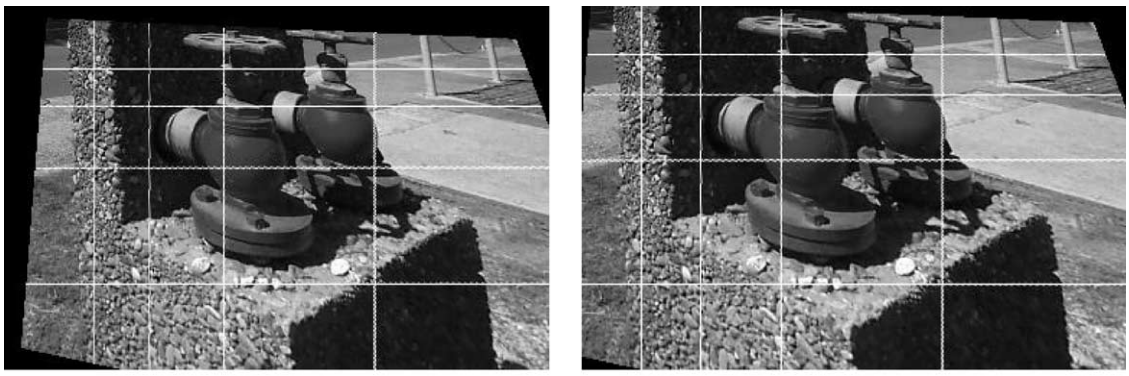
where $\mathbf{a} = (a_1, a_2, a_3)^T$ and

$$\mathbf{C}_2 = \begin{pmatrix} a_1 + 1 & a_2 & a_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Using the method described in [1], the a_1, a_2 and a_3 values in $\mathbf{C}_2$ can be estimated using least-squares technique. Therefore $\mathbf{P}_2$ can be calculated from $\mathbf{C}_2$ $\mathbf{P}_1$ $\mathbf{M}_{12}$. Matrix $\mathbf{P}_2$ will transform the epipole $\mathbf{e}_{21}$ in view 2 into the infinity. Therefore the original epipolar lines in view 2 can be made parallel to the image x-axes. The corresponding epipolar lines after transformation between view 1 and view 2 also lie on the same horizontal scanline.

Similarly the transformation matrix $\mathbf{P}_3$ for view 3 can be obtained from $\mathbf{F}_{13}$ and matching points between view 1 and view 3 as:

$$\begin{aligned}\mathbf{P}_3 &= (\mathbf{I} + \mathbf{P}_1\mathbf{e}_{13}\mathbf{b}^T)\mathbf{P}_1\mathbf{M}_{13} \\ &= (\mathbf{I} + (0\ 1\ 0)^T\mathbf{b}^T)\mathbf{P}_1\mathbf{M}_{13} = \mathbf{C}_3\mathbf{P}_1\mathbf{M}_{13}\end{aligned}$$



(a) (b)

Fig. 4. (a) Rectified reference image without the optimisation of the u, v values. (b) Rectified reference image with the optimisation of the u, v values.

where $\mathbf{b} = (b_1, b_2, b_3)^T$ and

$$\mathbf{C}_3 = \begin{pmatrix} 1 & 0 & 0 \\ b_1 & b_2 + 1 & b_3 \\ 0 & 0 & 1 \end{pmatrix}.$$

The values for b_1, b_2 and b_3 can be obtained similarly as a_1, a_2 and a_3. Matrix $\mathbf{P}_3$ will transform the epipole $\mathbf{e}_{31}$ in view 3 into the infinity. Therefore the original epipolar lines in view 3 can be made parallel to the image y-axes. The corresponding epipolar lines between view 1 and view 3 lie on the same vertical scanline.

Using the three transformation matrices $\mathbf{P}_1$, $\mathbf{P}_2$ and $\mathbf{P}_3$, the three original images can be rectified. When resampling the input images for rectification, bilinear interpolation can be used.

The process of obtaining **a** and **b** using least-squares technique involves minimising the distance of the matching points in a pair of views. The results of this is that the average disparities for the rectified images are minimised. However, this process may have the effect that some of the image regions are cut off in the rectified image. To reduce this effect, we can carry out a shift search similar to those mentioned in Section 2.2.4. The search needs only to be carried out in the horizontal direction (the u value) for view 2 and in the vertical direction (the v value) for view 3. The objective function of the search or optimisation is again the sum of distances for the four corners of the rectified and the original images. The effect of introducing of this search can be shown in Fig. 5.

After the trinocular image rectification, fast stereo matching algorithms can be used to derive disparities from the rectified images [22,9].

2.4. Algorithm steps

The steps of our algorithm for the trinocular image rectification are the following:

1 Automated sub-pixel features matching is carried out for the image triplet.

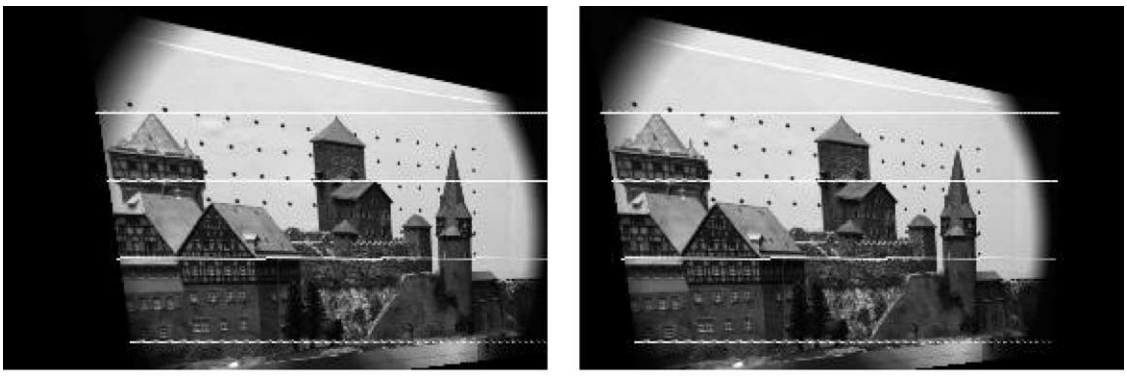
2 Obtain fundamental matrices. The fundamental matrices can either be obtained through the trilinear tensor or through projective invariants or estimated directly from image matching points. Here we list the main steps through the trilinear tensor method:
   a Image coordinate scaling and translation such that the centroid of all given matching coordinates is at the origin, and the average distance of a matching point from the origin is $\sqrt{2}$.
   b Solve a set of homogeneous equations using singular value decomposition (SVD) to obtain the trilinear tensor.
   c Calculate the fundamental matrices $\mathbf{F}_{12}$ and $\mathbf{F}_{13}$ from the trilinear tensor.

3 Uncalibrated trinocular image rectification:
   a Construct the transformation matrix $\mathbf{P}_1$ for the reference image using one of the methods described in Section 2.2.
   b Compute the other two transformation matrices $\mathbf{P}_2$ and $\mathbf{P}_3$ as described in Section 2.3.
   c Apply the three transformation matrices $\mathbf{P}_1$, $\mathbf{P}_2$ and $\mathbf{P}_3$ to the three images in view 1, view 2 and view 3 respectively, to obtain the rectified images.

3. Experimental results

This section shows some of the rectification results obtained using our new methods described in previous sections. A variety of real images have been tested.

For each triplet of images we have the left-right pair of images and left-top pair of images. We can use the method described in Section 2.1 for estimating the trilinear tensor or estimating the invariants in three images using six points and then calculating the fundamental matrices for the pairs of images in the triplet. We can also use Zhang et al's method for estimating the fundamental matrices directly for each pair of images [15]. After we have obtained the fundamental matrices for the left-right and the left-top pairs of images, we can then carry out the rectification process using our trinocular rectification algorithms.



(a) (b)

Fig. 5. (a) Rectified right image without the optimisation of the u value. (b) Rectified right image with the optimisation of the u value.

Fig. 6 gives an image triplet (left/reference in (b), right in (c) and top in (a) with some epipolar lines overlaid. In the reference image, there are two sets epipolar lines passing through two epipoles. Fig. 7 shows a sequence of transformed images using the vanishing points rectification method. Fig. 7(a) is the initial epipolar lines (the same as shown in Fig. 6(b). Fig 7(b) is the image obtained after applying the **PT** transformation. After this transformation, the epipolar lines in each set become parallel to each other. But the epipolar lines in different sets are not orthogonal to each other. Fig. 7(c) is the result after a skew transformation which makes the epipolar lines in the two set orthogonal to each other. Fig. 7(d) shows the final rectification result after an rotation transformation which aligns the two sets of epipolar lines to the image axes for the reference image. The rectification results for the image triplet are given in Fig. 8. Fig. 8(b) is the rectified reference image with two sets of epipolar lines orthogonal to each other and parallel to the image axes. Fig. 8(c) is the rectified image for the right image with horizontal epipolar lines. Fig. 8(a) is the rectified image for the top image with vertical epipolar lines.

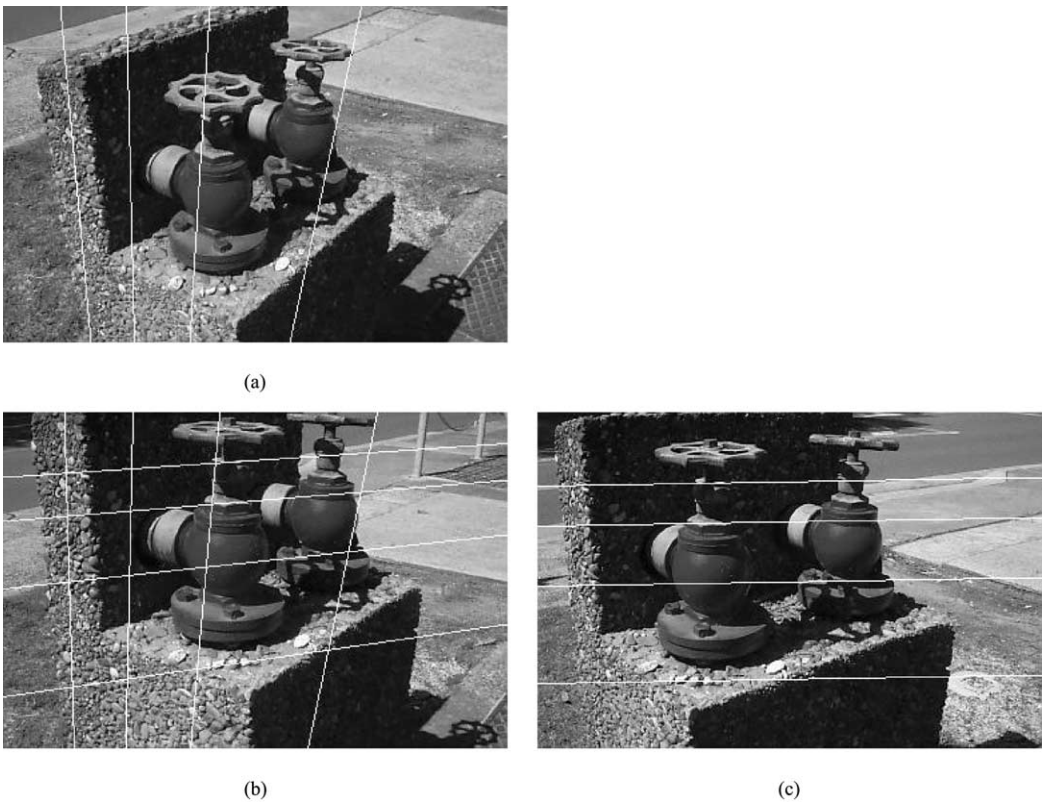
Fig. 9 shows a triplet of real images from the Calibrated Imaging Lab (CIL) Stereo dataset from CMU with epipolar lines overlaid. Note that although the images come with calibration information, the calibration information is not used for our uncalibrated image rectification tests. The lower-left image is the reference image which contains both near horizontal and near vertical epipolar lines. The top image has near vertical epipolar line; and the right image has near horizontal epipolar lines. The top-right drawing shows the relative camera positions for the three images. Fig. 10 shows the rectified images using our new methods. The epipolar lines are now either horizontal or vertical. Matching epipolar lines lie on the same horizontal or vertical scanlines.

Fig. 11 shows another triplet of images with epipolar lines before rectification. Fig. 12 gives the rectified images.

The amount of changes before and after image transformation during the rectification process depends upon the camera positions when taking the three images. For example if the third image is far from the standard 'top' position relative to the reference one, then the amount of image changes for this top image will be large.

Experiments have shown that the three methods for obtaining the projection matrix for the reference image described in Section 2.2 and the method for estimating the projection matrix for the other two views generate the same rectified images.

The rectification process (including the projection matrices estimation and image resampling) takes about 1.6 s with u,v optimisation and about 0.8 s without u,v optimisation for 756×504 pixel images on a 533 MHz DEC Alpha.



(a)

(b) (c)

Fig. 6. Epipolar lines overlaid on the original trinocular images.

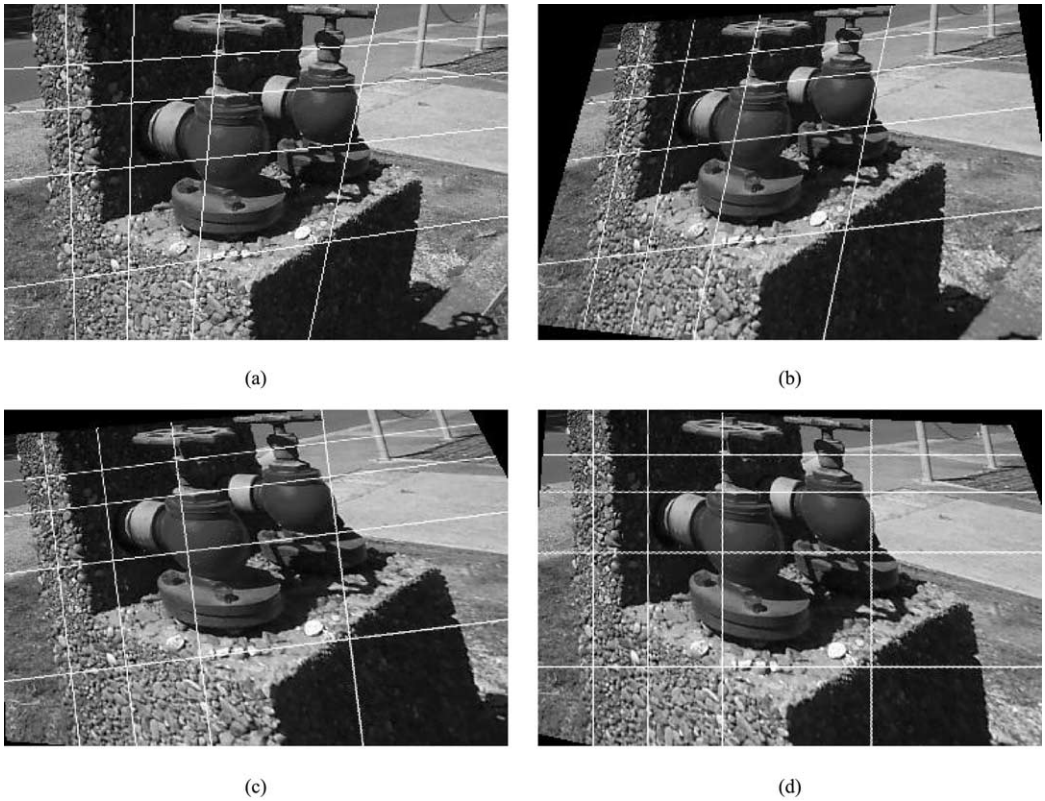


Fig. 7. Rectification steps for the reference image using the vanishing point method. (a) Original image with epipolar lines overlaid. (b) Image obtained after the translation and projection transformations. (c) Image obtained after skew operation on image in (b). (d) Rectified reference image after a further rotation operation. Epipolar lines are becoming horizontal or vertical.

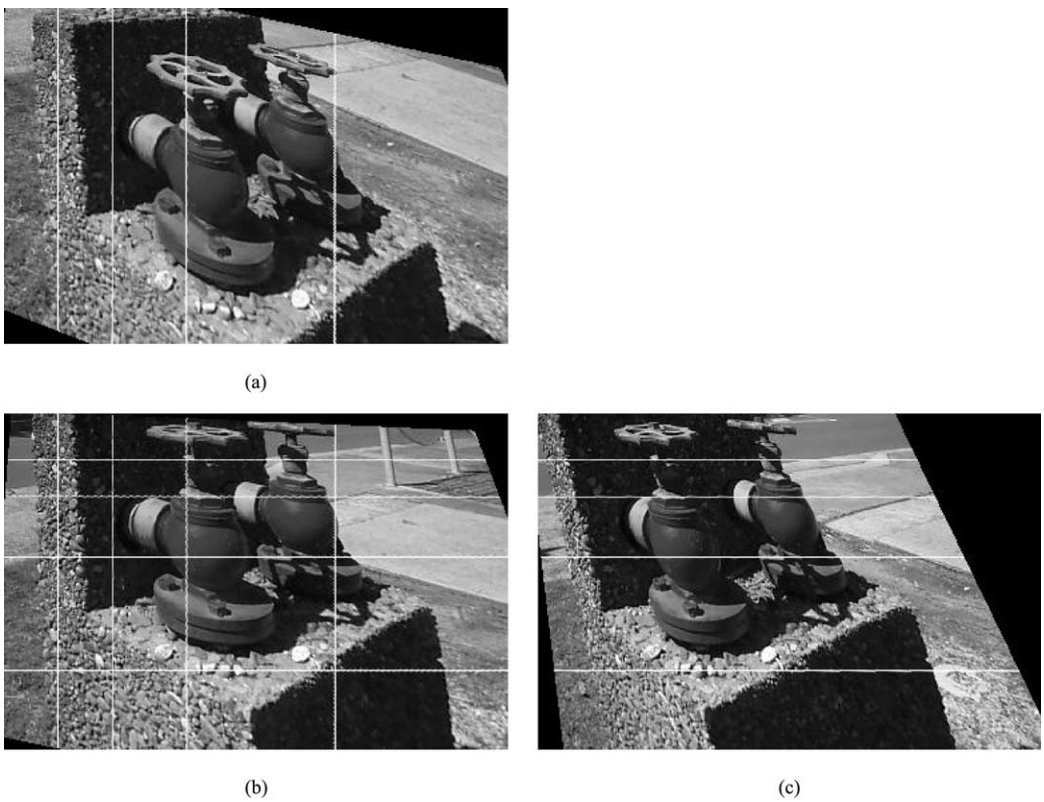


Fig. 8. Rectified trinocular images for those shown in Fig. 6. Epipolar lines are becoming horizontal (left-right) or vertical (left-top).

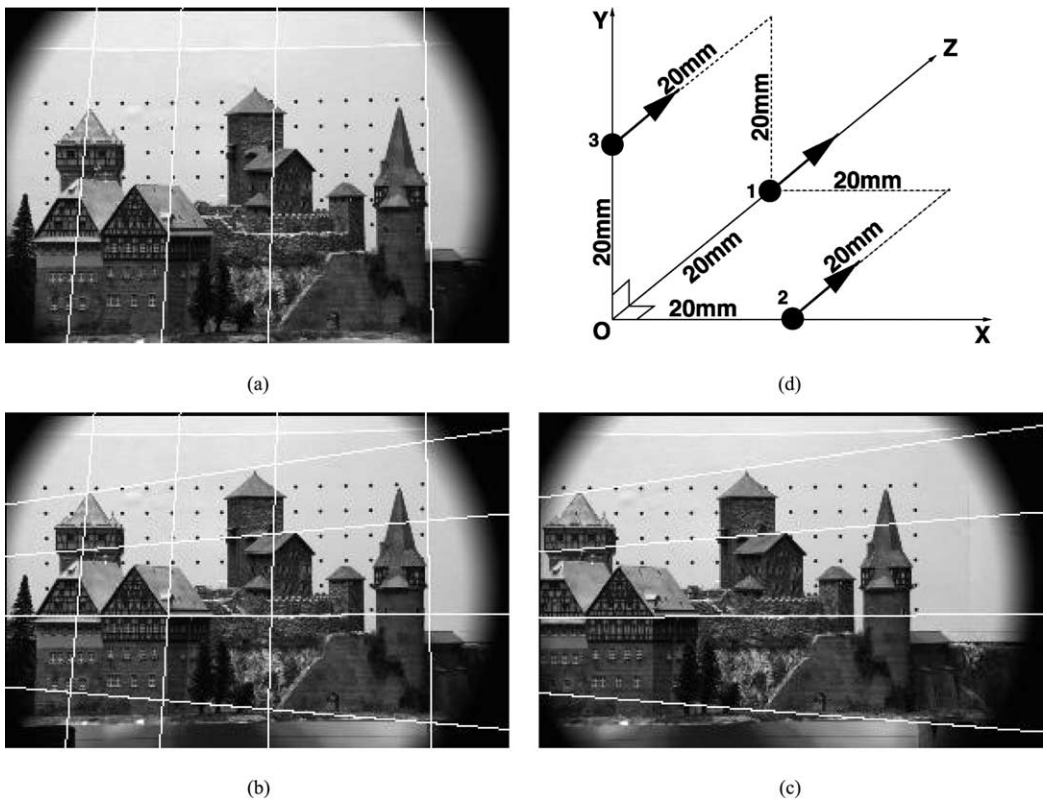


Fig. 9. Original trinocular images and imaging positions. (a) The top image with epipolar lines overlaid; (b) reference image with epipolar lines overlaid. Two sets of epipolar lines in the horizontal direction and vertical direction are shown; (c) the right image with epipolar lines overlaid; (d) illustration of the imaging geometry for this example. The information shown in the drawing (d) is not used during the rectification process. (CIL stereo images from the Calibrated Imaging Laboratory of the Carnegie Mellon University.)

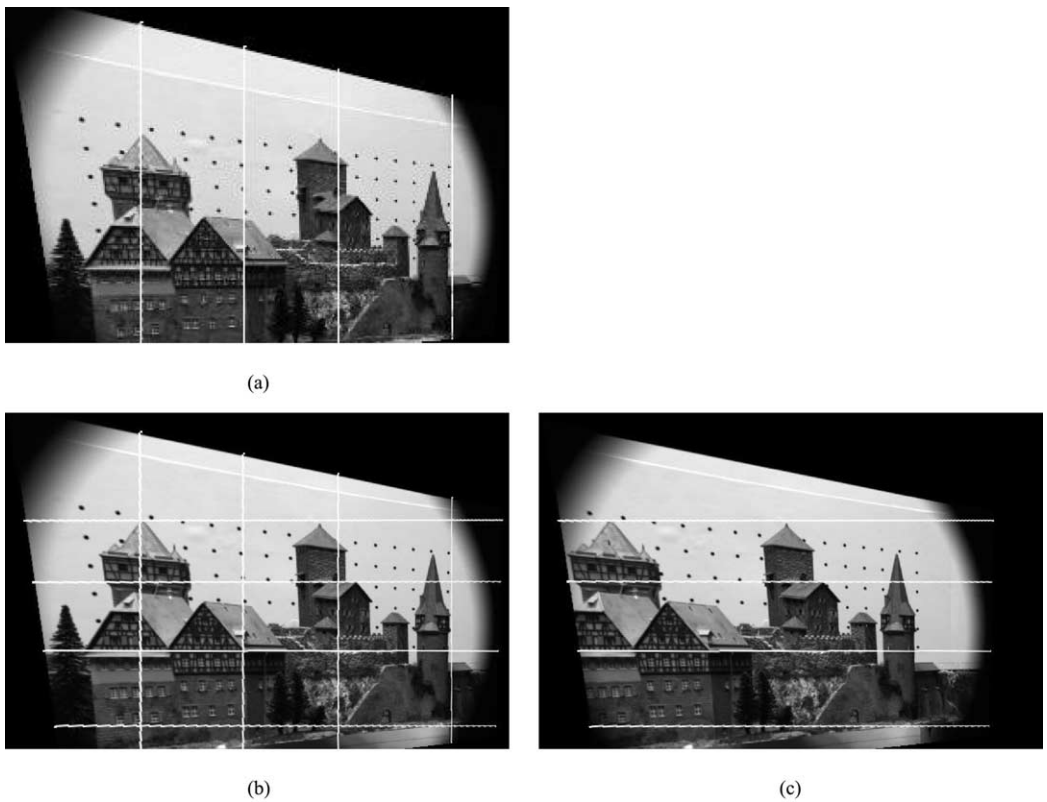
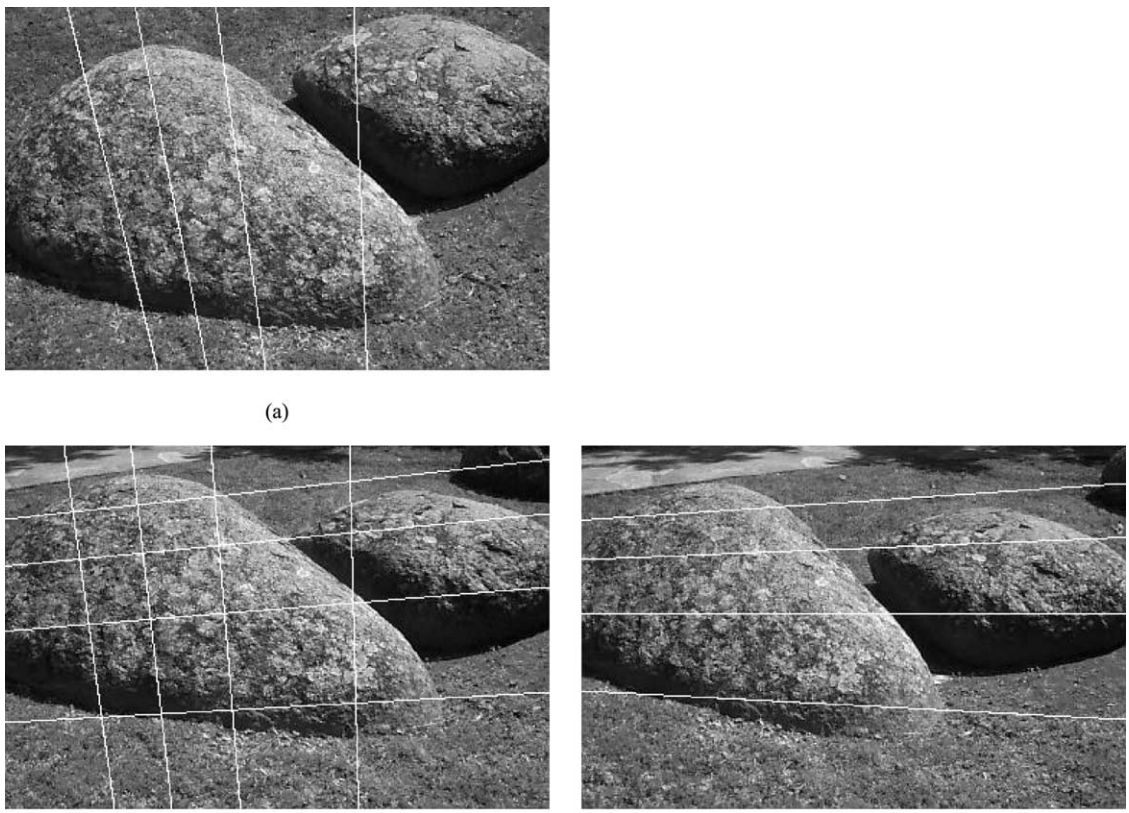


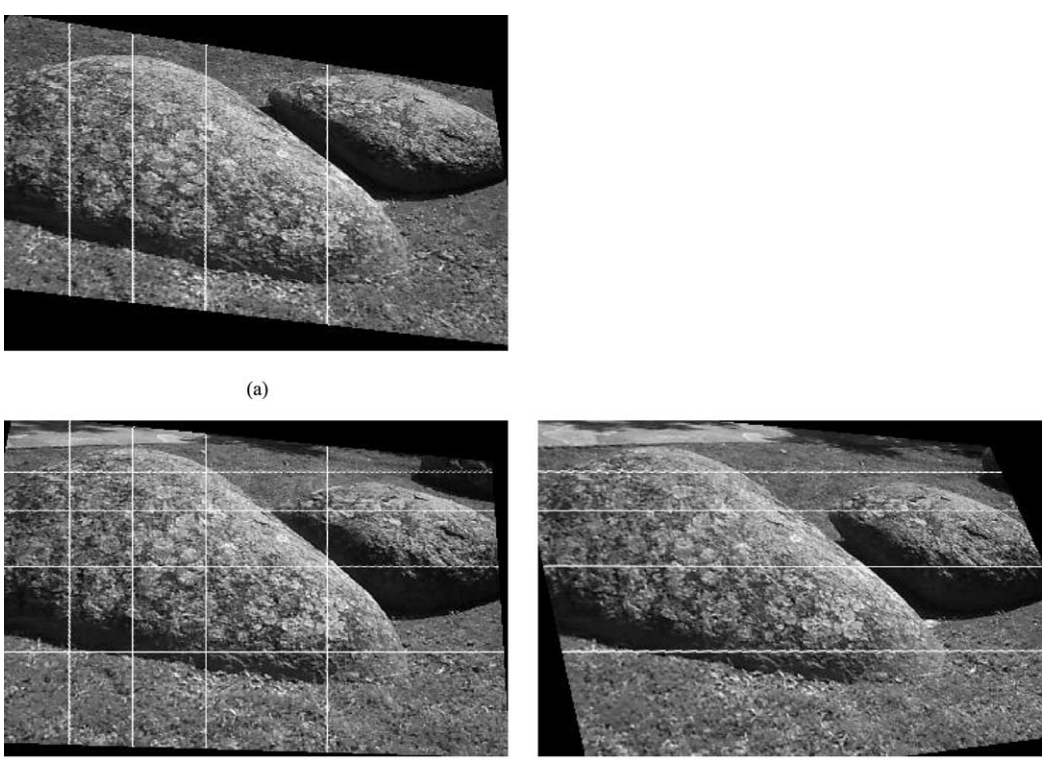
Fig. 10. Rectified trinocular images for those shown in Fig. 9. Epipolar lines are becoming horizontal (left-right) or vertical (left-top).



(a)

(b) (c)

Fig. 11. Epipolar lines overlaid on the original trinocular images.



(a)

(b) (c)

Fig. 12. Rectified trinocular images for those shown in Fig. 11. Epipolar lines are becoming horizontal (left-right) or vertical (left-top).

4. Conclusions

In this paper, several novel methods for rectifying uncalibrated trinocular images has been presented. The transformation matrices applied to the original images are constructed based on the epipolar geometries among the images triplet. Three novel methods have been proposed for estimating the rectification matrix for the reference image including the rotation and skew method, the affine transformation method, and the vanishing points method. Real images have been tested and the results validate the new methods.

5. Acknowledgement

The author thanks Dr Peter Corke of CSIRO and the anonymous referees for their comments and suggestions.

References

[1] R.I. Hartley, Theory and practice of projective rectification, International Journal of Computer Vision 35 (2) (1999) 115–127. November.

[2] 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. March.

[3] F. Isgro, E. Trucco, On robust rectification for uncalibrated images, In International Conference on Image Analysis and Processing, Venice, Italy. (1999) 297–302. 27-29 September, IEEE..

[4] M. Pollefeys, R. Koch, L. Van Gool, A simple and efficient rectification method for general motion, In Proceedings of International Conference on Computer Vision,Corfu, Greece. (1999) 496–501.

[5] C. Loop, Z. Zhang. Computing rectifying homographies for stereo vision. Technical Report MSR-TR-99-21, Microsoft Research, 8 April 1999.

[6] D. Papadimitriou, T. Dennis, Epipolar line estimation and rectification for stereo image pairs, IEEE Transactions on Image Processing 5 (4) (1996) 672–676.

[7] N. Ayache, C. Hansen, Rectification of images for binocular and trinocular stereovision, In Proceedings of International Conference on Pattern Recognition, Ergif Palace Hotel, Rome, Italy vol. 1 (1988) 11–16. November.

[8] J. Shao, C. Fraser, Rectification and matching of trinocular imagery, Geomatics Research Australasia 71 (1999) 73–86.

[9] http://www.ptgrey.com/.

[10] A. Fusiello, E. Trucco, A. Verri, A. compact, algorithm for rectification of stereo pairs, Machine Vision Applications 12 (1) (2000) 16–22.

[11] A. Shashua, Algebraic functions for recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (8) (1995) 779–789.

[12] R.I. Hartley, In defence of the 8-point algorithm, In Proceedings of International Conference on Computer Vision, MIT,Cambridge, Massachusetts,IEEE 20-23 (1995) 1064–1070. June.

[13] A. Shashua, M. Werman, Trilinearity of three perspective views and its associated tensor, In Proceedings of International Conference on Computer Vision,Boston. (1995) 920–925. June.

[14] L. Quan, Invariants of six points and projective reconstruction from three uncalibrated images, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (1) (1995) 34–46.

[15] Z. Zhang, R. Deriche, O. Faugeras, Q.-T. Luong, A. robust, technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry, Artificial Intelligence 78 (1995) 87–119.

[16] P.H.S. Torr, A. Zisserman, Robust parameterization and computation of the trifocal tensor, Image and Vision Computing 15 (8) (1997) 591–605. August.

[17] Andrew Zisserman Richard Hartley. Multiple View Geometry in Computer Vision. Cambridge University Press, 2000.

[18] Robert T. Collins, J. Ross Beveridge. Matching perspective views of coplanar structures using projective unwarping and similarity matching. Technical Report CMPSCI TR94-06, Dept. of Computer Science, University of Massachusetts, Amherst, February 1994.

[19] A. Criminisi, A. Zisserman, Shape from texture: homogeneity revisited, In Proceedings of 11th British Machine Vision Conference, Bristol, UK. (2000) 82–91. September.

[20] J.H. Im, Y.I. Yoon, J.H. Kim, S.H. Kim, J.S. Choi, Plane rectification using bilinear forms on projective spaces, In FCV: 7th Korea-Japan Joint Workshop on Computer Vision (2001) 19–24. February.

[21] D. Liebowitz, A. Zisserman, Metric rectification for perspective images of planes, In Proceedings of Computer Vision and Pattern Recognition (1998) 482–488.

[22] C. Sun, Fast stereo matching using rectangular subregioning and 3D maximum-surface techniques, International Journal of Computer Vision 47 (1/2/3) (2002) 99–117. April-June.