TRINOCULAR RECTIFICATION FOR VARIOUS CAMERA SETUPS

Matthias Heinrichs* and Volker Rodehorst

Computer Vision & Remote Sensing, Berlin University of Technology, Franklinstr. 28/29, FR 3-1,
D-10587 Berlin, Germany – (matzeh, vr)@cs.tu-berlin.de

KEY WORDS: Photogrammetry, Trifocal Geometry, Planar Rectification, Uncalibrated Images, Normal Images

ABSTRACT:

Convergent stereo images can be rectified such that they correspond to the stereo normal case. The important advantage is that the correspondence analysis is simplified to a linear search along one of the image axes. The proposed method describes a rectification method for three uncalibrated cameras, where the configuration is not generally known. The algorithm automatically determines the best positioning of the cameras and corrects mirroring effects to fit the desired camera setup. Since there is no direct solution for the rectification problem for three cameras, the rectification homographies are linearly determined to within six degrees of freedom from three compatible fundamental matrices. The remaining six parameters are then obtained by enforcing additional constraints.

1. INTRODUCTION

The epipolar geometry of a pinhole camera model implies that the correspondent of a given point lies on its epipolar line in the second image of a stereo pair. The pencil of all epipolar lines passes through a point called the epipole **e**, which is the projection of the camera center onto the corresponding image plane. In case of convergent stereo setups, the image matching task is fairly complex and thus inefficient. Rectification determines a transformation of each image plane such that the epipolar lines become parallel to one of the image axes (see Figure 1). This configuration corresponds to the stereo normal case.

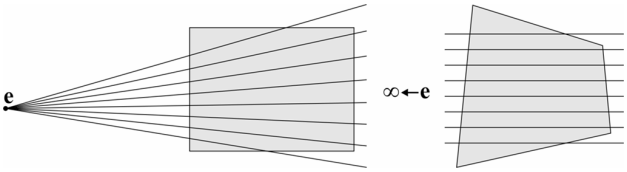

Figure 1. Geometry of normal images

A discussion of *binocular rectification* methods can be found in Hartley (Hartley, 1999), where the normal images are generated using a single linear transformation, which is often referred to as *planar rectification*. In principle, the images are reprojected onto a plane parallel to the baseline between the optical centers. This technique is relatively simple, fast and retains image characteristics, i.e. straight lines. The reported methods minimize image distortion and maximize the computational efficiency. Matoušek (Matoušek et al., 2004) proposed a data-optimal rectification procedure that minimizes the loss of discriminability in the resulting images.

The linear approach however is not general and fails when the epipole lies within the image. The transformation for an epipole close to the image borders leads to an extremely large and strongly distorted image. This problem can be avoided by using stereo configurations with almost parallel camera alignment. For general image sequences with arbitrary camera orientations, an improved method must be used.

In order to solve this problem, Roy (Roy et al., 1997) suggested a *cylindrical rectification* method with a separate transformation for each epipolar line. The basic idea lies in the use of polar coordinates with the epipole at the origin. Pollefeys (Pollefeys et al., 1999) adapted this non-linear approach for applications in projective geometry. However, the use of different non-linear transformations leads to irregular, distorted images, which makes the following correspondence analysis more difficult. To avoid this effect, a *hybrid* procedure was proposed by Oram (Oram, 2001). Here, the epipoles are first overlaid with a compatible homography and after which the same non-linear transformation is used for both images.

It has been shown that multi-view matching can considerably improve the quality of spatial reconstruction, in which rectification remains of interest. In case of a *trinocular rectification*, the images are reprojected onto a plane, which lies parallel to the optical centers (see Figure 2).

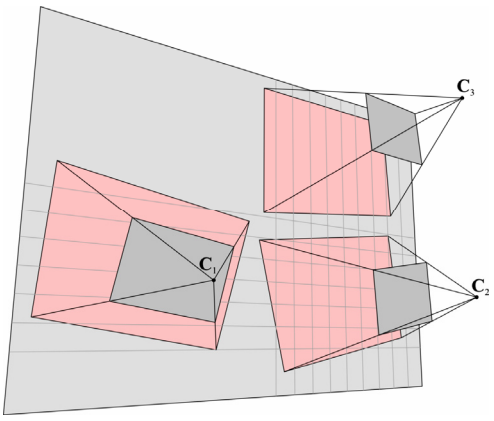

Figure 2. Trinocular Rectification

Ayache and Hansen (Ayache et. al., 1988) proposed a technique for rectifying image triplets that works with calibrated cameras. Loop and Zhang (Loop and Zhang, 1999) presented a stratified method to decompose each transformation and formulate geometric criteria to minimize image distortion during rectification. For a review of various trinocular rectification methods, see (Sun, 2003).

Zhang (Zhang et al., 2002/03) proposed rectification homographies in a closed form and introduced stronger, geometrically meaningful constraints. An (An et al., 2004) reported an efficient trinocular rectification method using the geometric camera model instead of the relative image orientation. However, this method is only applicable when well known control points are available to calibrate and orientate the cameras.

The proposed trinocular rectification method requires an uncalibrated image triple with more or less parallel camera alignment. The camera configuration is arbitrary, but each projection center must be invisible in all other images. This condition is necessary, since otherwise the epipoles lie in the image and mapping them to infinity will lead to unacceptable distortion of the images. Furthermore, we assume non-degenerate camera positions, where the camera centers are not collinear, because collinear setups can be rectified by chaining a classical binocular rectification approach. Additionally, a common overlapping area and at least six homologous image points are necessary, so that the trifocal tensor, the fundamental matrices and the epipoles can be determined (Hartley and Zisserman, 2000). The result consists of three geometrically transformed images, in which the epipolar lines run parallel to the image axes.

2. CAMERA SETUP

A given image triplet consists of the original images *b* (base), *h* (horizontal) and *v* (vertical). Subsequently, we denote the rectified images $\tilde{b}$, $\tilde{h}$ and $\tilde{v}$. The rectification tries to fit any image triple to a configuration shown in Figure 3.

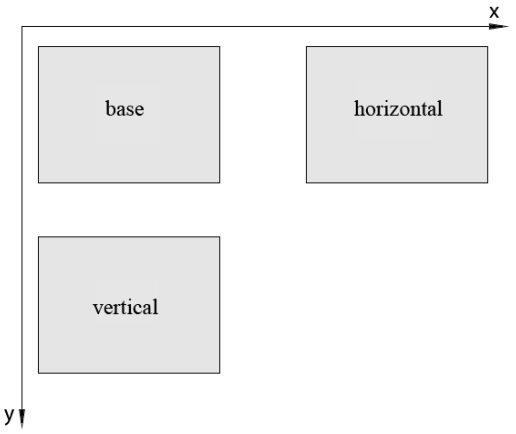

Figure 3. Image arrangement

This setup has the following properties:

- The epipolar lines of image *b* and image *h* correspond with their image rows.
- The epipolar lines of image *b* and image *v* correspond with their image columns.
- The epipolar lines of image *h* and image *v* have a slope of minus unity.

The last property has the advantage, that the disparities between corresponding points in $\tilde{b} \leftrightarrow \tilde{h}$ and $\tilde{b} \leftrightarrow \tilde{v}$ are equal.
For rectification, the epipoles between the images *b*, *h* and *v* should be mapped to infinity:

$$\begin{aligned} \tilde{\mathbf{e}}_{bh} &= \tilde{\mathbf{e}}_{hb} = [1 \quad 0 \quad 0]^{\mathrm{T}} \\ \tilde{\mathbf{e}}_{bv} &= \tilde{\mathbf{e}}_{vb} = [0 \quad 1 \quad 0]^{\mathrm{T}} \\ \tilde{\mathbf{e}}_{hv} &= \tilde{\mathbf{e}}_{vh} = [-1 \quad 1 \quad 0]^{\mathrm{T}} \end{aligned} \tag{1}$$

The relative image orientation for this setup is quite simple. The fundamental matrices between the rectified images are given by

$$\tilde{\mathbf{F}}_{bh} = [\tilde{\mathbf{e}}_{bh}]_{\times} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$\tilde{\mathbf{F}}_{bv} = [\tilde{\mathbf{e}}_{bv}]_{\times} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \text{ and} \tag{2}$$

$$\tilde{\mathbf{F}}_{hv} = [\tilde{\mathbf{e}}_{hv}]_{\times} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}.$$

2.1 Computation of the Projective Image Orientation

This section describes a method to assign each image in a given triplet to a suitable position *b*, *h* and *v*. At least six homologous image points are used to compute the trifocal tensor $\mathcal{T}$ by the minimal 6-point-algorithm (Torr and Zisserman, 1997) with the robust estimator GASAC (Rodehorst and Hellwich, 2006). The projection matrix $\mathbf{P}_1$ of the reference camera is set to a canonical form and the projection matrices $\mathbf{P}_2$ and $\mathbf{P}_3$ for the other two cameras can be obtained from $\mathcal{T}$. The projection center $\mathbf{C}_1 = (C_{11}, C_{12}, C_{13}, C_{14})^{\mathrm{T}}$ of the first camera is placed in the origin and the two remaining projection centers $\mathbf{C}_2$ and $\mathbf{C}_3$ can be estimated from the projective $\mathbf{P}_2$ and $\mathbf{P}_3$.
$\mathbf{P}_1$ is defined to point in z-direction and all cameras have a common overlapping area and are not visible in the other cameras. Therefore $\mathbf{P}_2$ and $\mathbf{P}_3$ can not have a significant translation in z-direction and the camera alignment can be computed in the x/y-plane. Under these conditions the absolute angles β_1, β_2 and β_3 between the cameras in the x/y-plane can be computed:

$$\begin{aligned} \beta_{1_1} &= \begin{cases} \left|\arctan\left(C_{22} / C_{21}\right)\right| & \text{for } C_{21} \neq 0 \\ \pi / 2 & \text{otherwise} \end{cases} \\ \beta_2 &= \begin{cases} \left|\arctan\left(C_{32} / C_{31}\right)\right| & \text{for } C_{31} \neq 0 \\ \pi / 2 & \text{otherwise} \end{cases} \\ \beta_3 &= \begin{cases} \left|\arctan\left(\left(C_{32} - C_{22}\right) / \left(C_{31} - C_{21}\right)\right)\right| & \text{for} \left(C_{31} - C_{21}\right) \neq 0 \\ \pi / 2 & \text{otherwise} \end{cases} \end{aligned} \tag{3}$$

Using the absolute value is necessary, since camera configurations, which are different from the setup assumed in Figure 3 produce mirror effects. The compensation of these effects is discussed in section 3.3.3. Figure 3 shows that the camera pair with the highest angle value is the pair $b \leftrightarrow v$ and the camera pair with the lowest angle value is the pair $b \leftrightarrow h$. Now the images can be aligned in a suitable fashion and the trifocal tensor $\mathcal{T}$ must be adapted for this enhanced setup.

3. RECTIFICATION

The initial task is to determine the relative image orientation. The fundamental matrices of the original images can be obtained uniquely by the trifocal tensor (see section 2.1). Note that the fundamental matrices are not independent and have only 18 significant parameters in total (Hartley and Zisserman, 2000):

$$\mathbf{e}_{hv}^{\mathrm{T}} \mathbf{F}_{hb} \mathbf{e}_{bv} = \mathbf{e}_{vb}^{\mathrm{T}} \mathbf{F}_{vh} \mathbf{e}_{hb} = \mathbf{e}_{vh}^{\mathrm{T}} \mathbf{F}_{vb} \mathbf{e}_{bh} = 0$$

3.1 Mapping Epipoles to Infinity

Let $\mathbf{H}_b$, $\mathbf{H}_h$ and $\mathbf{H}_v$ be the unknown homographies between the original and rectified images. The rows of these homographies will be abbreviated by three vectors **u**, **v** and **w**:

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{u}_i^{\mathrm{T}} \\ \mathbf{v}_i^{\mathrm{T}} \\ \mathbf{w}_i^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} u_{i1} & u_{i2} & u_{i3} \\ v_{i1} & v_{i2} & v_{i3} \\ w_{i1} & w_{i2} & w_{i3} \end{bmatrix} \quad \text{for} \quad i \in \{b,h,v\} \tag{4}$$

For a correspondence $\mathbf{x}_b \leftrightarrow \mathbf{x}_h \leftrightarrow \mathbf{x}_v$, the projective transformation between the image coordinates can be written as

$$\tilde{\mathbf{x}}_i = \mathbf{H}_i \mathbf{x}_i \quad \text{for} \quad i \in \{b,h,v\}. \tag{5}$$

The fundamental matrices, which are calculated from the original images, satisfy the epipolar constraints:

$$\begin{aligned} \mathbf{x}_h^{\mathrm{T}} \mathbf{F}_{bh} \mathbf{x}_b &= 0 \\ \mathbf{x}_v^{\mathrm{T}} \mathbf{F}_{bv} \mathbf{x}_b &= 0 \\ \mathbf{x}_v^{\mathrm{T}} \mathbf{F}_{hv} \mathbf{x}_h &= 0 \end{aligned} \tag{6}$$

Similar conditions apply for the rectified images:

$$\begin{aligned} \tilde{\mathbf{x}}_h^{\mathrm{T}} \tilde{\mathbf{F}}_{bh} \tilde{\mathbf{x}}_b &= 0 \\ \tilde{\mathbf{x}}_v^{\mathrm{T}} \tilde{\mathbf{F}}_{bv} \tilde{\mathbf{x}}_b &= 0 \\ \tilde{\mathbf{x}}_v^{\mathrm{T}} \tilde{\mathbf{F}}_{hv} \tilde{\mathbf{x}}_h &= 0 \end{aligned} \tag{7}$$

Combining equations (6) and (7), one obtains

$$\begin{aligned} \mathbf{x}_h^{\mathrm{T}} \mathbf{H}_h^{\mathrm{T}} \tilde{\mathbf{F}}_{bh} \mathbf{H}_b \mathbf{x}_b &= \mathbf{x}_h^{\mathrm{T}} \mathbf{F}_{bh} \mathbf{x}_b = 0 \\ \mathbf{x}_v^{\mathrm{T}} \mathbf{H}_v^{\mathrm{T}} \tilde{\mathbf{F}}_{bv} \mathbf{H}_b \mathbf{x}_b &= \mathbf{x}_v^{\mathrm{T}} \mathbf{F}_{bv} \mathbf{x}_b = 0 \\ \mathbf{x}_v^{\mathrm{T}} \mathbf{H}_v^{\mathrm{T}} \tilde{\mathbf{F}}_{hv} \mathbf{H}_h \mathbf{x}_h &= \mathbf{x}_v^{\mathrm{T}} \mathbf{F}_{hv} \mathbf{x}_h = 0 \end{aligned} \tag{8}$$

and comparing the result with (7), it follows that

$$\begin{aligned} \mathbf{H}_h^{\mathrm{T}} \tilde{\mathbf{F}}_{bh} \mathbf{H}_b &= \lambda_1 \mathbf{F}_{bh} \\ \mathbf{H}_v^{\mathrm{T}} \tilde{\mathbf{F}}_{bv} \mathbf{H}_b &= \lambda_2 \mathbf{F}_{bv} \\ \mathbf{H}_v^{\mathrm{T}} \tilde{\mathbf{F}}_{hv} \mathbf{H}_h &= \lambda_3 \mathbf{F}_{hv} \end{aligned} \tag{9}$$

where λ_i are scale factors. The rectified fundamental matrices $\tilde{\mathbf{F}}_{bh}$, $\tilde{\mathbf{F}}_{bv}$ and $\tilde{\mathbf{F}}_{hv}$ contain many zeros (see Eq. 2). Hence, equations (9) can be simplified to give:

$$\begin{aligned} \mathbf{w}_h \mathbf{v}_b{}^{\mathrm{T}} - \mathbf{v}_h \mathbf{w}_b{}^{\mathrm{T}} &= \lambda_1 \mathbf{F}_{bh} \\ \mathbf{u}_v \mathbf{w}_b{}^{\mathrm{T}} - \mathbf{w}_v \mathbf{u}_b{}^{\mathrm{T}} &= \lambda_2 \mathbf{F}_{bv} \\ (\mathbf{u}_v + \mathbf{v}_v) \mathbf{w}_h{}^{\mathrm{T}} - \mathbf{w}_v (\mathbf{u}_h + \mathbf{v}_h)^{\mathrm{T}} &= \lambda_3 \mathbf{F}_{hv} \end{aligned} \tag{10}$$

3.2 Computation of Rectifying Homographies

Since the fundamental matrices are of rank 2, the equations (10) can not be solved directly. However, knowing the vectors $\mathbf{w}_b$, $\mathbf{w}_h$ and $\mathbf{w}_v$, gives a solution for (10) with six degrees of freedom (DOF). The **w**-vectors have a convenient property: Since the epipoles in the original images should be mapped to infinity, the scalar product of $\mathbf{w}_i$ with both epipoles of an image is zero. That means the **w**-vector is perpendicular to both epipoles of an image and can be calculated (up to a scale factor) by the cross product of the two epipoles:

$$\begin{aligned} \mathbf{w}_b &= \mathbf{e}_{bh} \times \mathbf{e}_{bv} \\ \mathbf{w}_h &= \mathbf{e}_{hb} \times \mathbf{e}_{hv} \\ \mathbf{w}_v &= \mathbf{e}_{vb} \times \mathbf{e}_{vh} \end{aligned} \tag{11}$$

The epipoles are the left and right null-vectors of the **F**-matrices and can be determined by singular value decomposition (Hartley and Zisserman, 2000). Since the projection centers are not collinear, all epipole pairs are linearly independent. Hence their cross products, especially the third components, are non zero and the third component of each **w**-vector can be scaled to unity to simplify the equations (10).

Six variables have to be defined for a direct solution. Depending on the variables chosen, the equations become very simple. We recommend setting

$$u_{b3} = v_{h3} = v_{v3} = 0 \quad \text{and} \quad \lambda_1 = \lambda_2 = \lambda_3 = 1 \tag{12}$$

which yields the solutions

$$\mathbf{H}_b^* = \begin{bmatrix} w_{b1} F_{33}^{bv} - F_{31}^{bv} & w_{b2} F_{33}^{bv} - F_{32}^{bv} & 0 \\ F_{31}^{bh} & F_{32}^{bh} & F_{33}^{bh} \\ w_{b1} & w_{b2} & 1 \end{bmatrix} \tag{13}$$

$$\mathbf{H}_h^* = \begin{bmatrix} w_{h1}(F_{33}^{bv} - F_{33}^{bh}) + F_{13}^{bh} - F_{31}^{hv} & w_{h2}(F_{33}^{bv} - F_{33}^{bh}) + F_{23}^{bh} - F_{32}^{hv} & F_{33}^{bv} - F_{33}^{hv} \\ w_{h1} F_{33}^{bh} - F_{13}^{bh} & w_{h2} F_{33}^{bh} - F_{23}^{bh} & 0 \\ w_{h1} & w_{h2} & 1 \end{bmatrix}$$

$$\mathbf{H}_v^* = \begin{bmatrix} F_{13}^{bv} & F_{23}^{bv} & F_{33}^{bv} \\ w_{v1}(F_{33}^{bv} - F_{33}^{hv}) + F_{13}^{hv} - F_{13}^{bv} & w_{v2}(F_{33}^{bv} - F_{33}^{hv}) + F_{23}^{hv} - F_{23}^{bv} & 0 \\ w_{v1} & w_{v2} & 1 \end{bmatrix}.$$

They look similar to those proposed by Zhang (Zhang et al., 2002), but satisfy the common image representation with the origin in the upper left corner. These equations already describe primitive rectifying homographies for the given image triplet, but generally produce undesirable shearing and scaling. A detailed derivation of (13) from (10) is available online (Heinrichs and Rodehorst, 2006).

3.3 Imposing Geometric Constraints

The assumptions in equation (12) can be generalized once more by comparing equations (13) and (10). The missing variables will be split into different components with the geometric meaning of translation, mirroring, scaling and shearing. The mirroring component is necessary because the computation of the camera setup ignores mirror effects in (3). These two parameters are introduced for convenience, to maintain the order of the image content, but do not influence the correctness of the rectification. To clarify the meaning of the remaining DOF, we define:

$$\begin{aligned} &s_1 = u_{b3}, \quad s_2 = v_{h3}, \quad s_3 = v_{v3} - v_{h3} \quad \text{and} \\ &\alpha_1 = \lambda_1 / \lambda_3, \quad \alpha_2 = \lambda_2 / \lambda_3, \quad \alpha_3 = \lambda_3 \end{aligned} \tag{14}$$

The mirroring components can be written as follows:

$$m_x, m_y \in \{-1, 1\} \tag{15}$$

The choice of correct signs allows a better visual interpretation. Mirror compensation is necessary for camera setups, in which the original images are flipped over one or two axes: *v* below *b* or *h* left of *b*. If m_x and m_y have different signs the slope of the epipolar lines between $\tilde{h}$ and $\tilde{v}$ becomes positive.

The general solution for the homographies is given by:

$$\mathbf{H}_b = \begin{bmatrix} 1 & 0 & s_1 \\ 0 & 1 & s_2 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} m_x \alpha_3 & 0 & 0 \\ 0 & m_y \alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \alpha_2 & 0 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{H}_b^* \tag{16}$$

$$\mathbf{H}_h = \begin{bmatrix} 1 & 0 & s_1 + s_3 \\ 0 & 1 & s_2 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} m_x \alpha_3 & 0 & 0 \\ 0 & m_y \alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1-\alpha_1 & F_{33}^{bv}(\alpha_2 - 1) \\ 0 & \alpha_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{H}_h^*$$

$$\mathbf{H}_v = \begin{bmatrix} 1 & 0 & s_1 \\ 0 & 1 & s_2 + s_3 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} m_x \alpha_3 & 0 & 0 \\ 0 & m_y \alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \alpha_2 & 0 & 0 \\ 1-\alpha_2 & 1 & F_{33}^{bv}(\alpha_2 - 1) \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{H}_v^*$$

where the free parameters can be interpreted as:

- s_1 is the global shift value in x-direction of all images
- s_2 defines the global shift value in y-direction
- s_3 is the shift value in x-direction for image h and the shift value in y-direction of image v relative to image b
- α_1 is the scale of the y-component of images b and h, which affects the shearing in y-direction of image v
- α_2 is the scale of the x-component of images b and v, which affects the shearing in x-direction of image h
- α_3 defines the global scaling factor to keep the images at a suitable resolution

The two convenience parameters are:

- m_x is a mirroring factor in the x-direction of all images
- m_y defines the mirroring factor in the y-direction

The challenge is to estimate optimal values for these parameters. Since the first six parameters are independent of each other, one can deal with them separately. The shift parameters depend on the mirroring parameters, thus we have to correct the mirror parameters first. The values should be calculated in the following order:

1. Finding proper shearing values for α_1 and α_2
2. Finding a global scale value α_3
3. Compensate potential mirroring using m_x and m_y
4. Finding right offset values for s_1, s_2 and s_3

The factor $F_{33}^{by}(\alpha_2 - 1)$ in the shearing matrices of equation (16) of image h and v is needed to compensate the loss of information in (10) by applying (12).

3.3.1 Shearing Correction

Following the approach of Loop and Zhang (Loop and Zhang, 1999), the shearing of images h and v can be minimized by keeping two perpendicular vectors in the middle of the original image perpendicular in the rectified one. Let W be the width and H be the height of the original images. Two perpendicular direction vectors $\mathbf{x}$ and $\mathbf{y}$ are defined by computing the center lines of the image.

$$\mathbf{x} = \mathbf{a} - \mathbf{b}, \quad \mathbf{y} = \mathbf{c} - \mathbf{d} \quad \text{with}$$

$$\mathbf{a} = \begin{bmatrix} W/2 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} W/2 \\ H \\ 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0 \\ H/2 \\ 1 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} W \\ H/2 \\ 1 \end{bmatrix} \tag{17}$$

For the horizontal rectified image $\tilde{h}$ the vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ can be calculated as follows:

$$\begin{gathered} \tilde{\mathbf{x}} = \tilde{\mathbf{a}} - \tilde{\mathbf{b}}, \quad \tilde{\mathbf{y}} = \tilde{\mathbf{c}} - \tilde{\mathbf{d}} \\ \tilde{\mathbf{a}} = \mathbf{H}_h^* \cdot \mathbf{a}, \quad \tilde{\mathbf{b}} = \mathbf{H}_h^* \cdot \mathbf{b}, \quad \tilde{\mathbf{c}} = \mathbf{H}_h^* \cdot \mathbf{c}, \quad \tilde{\mathbf{d}} = \mathbf{H}_h^* \cdot \mathbf{d} \end{gathered} \tag{18}$$

We apply the shearing matrix to $\mathbf{H}_h$, which can be derived from (16):

$$\mathbf{S}_h = \begin{bmatrix} 1 & 1-\alpha_1 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{19}$$

Note that the factor $F_{33}^{by}(\alpha_2 - 1)$ in (16) is ignored, since the last (homogeneous) component of the vectors $\tilde{\mathbf{a}}$, $\tilde{\mathbf{b}}$, $\tilde{\mathbf{c}}$ and $\tilde{\mathbf{d}}$ are the same and therefore their difference equals zero.
The vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ are perpendicular, when the constraint

$$(\mathbf{S}_h \tilde{\mathbf{x}})^{\mathrm{T}} (\mathbf{S}_h \tilde{\mathbf{y}}) = 0 \tag{20}$$

is satisfied. The quadratic equation

$$\alpha_1^2 + \alpha_1 \left(\frac{-\tilde{x}_x \tilde{y}_y - \tilde{x}_y \tilde{y}_x - 2\tilde{x}_y \tilde{y}_y}{2\tilde{x}_y \tilde{y}_y} \right) + \left(\frac{\tilde{x}_x \tilde{y}_x + \tilde{x}_x \tilde{y}_y + \tilde{x}_y \tilde{y}_x + \tilde{x}_y \tilde{y}_y}{2\tilde{x}_y \tilde{y}_y} \right) = 0 \tag{21}$$

has two solutions, where the solution with the smaller absolute value $|\alpha_1|$ is to be preferred.
For the vertical rectified image $\tilde{v}$, the vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ can be calculated as follows:

$$\begin{gathered} \tilde{\mathbf{x}} = \tilde{\mathbf{a}} - \tilde{\mathbf{b}}, \quad \tilde{\mathbf{y}} = \tilde{\mathbf{c}} - \tilde{\mathbf{d}} \\ \tilde{\mathbf{a}} = \mathbf{H}_v^* \cdot \mathbf{a}, \quad \tilde{\mathbf{b}} = \mathbf{H}_v^* \cdot \mathbf{b}, \quad \tilde{\mathbf{c}} = \mathbf{H}_v^* \cdot \mathbf{c}, \quad \tilde{\mathbf{d}} = \mathbf{H}_v^* \cdot \mathbf{d} \end{gathered} \tag{22}$$

The shearing matrix can be written as

$$\mathbf{S}_v = \begin{bmatrix} \alpha_2 & 0 & 0 \\ 1-\alpha_2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{23}$$

Once again, the factor $F_{33}^{by}(\alpha_2 - 1)$ can be ignored, because the third component of $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ is zero. The perpendicularity condition results in

$$(\mathbf{S}_v \tilde{\mathbf{x}})^{\mathrm{T}} (\mathbf{S}_v \tilde{\mathbf{y}}) = 0 \,. \tag{24}$$

Again, after solving the quadratic equation

$$\alpha_2^2 + \alpha_2 \left(\frac{-\tilde{x}_x \tilde{y}_y - \tilde{x}_y \tilde{y}_x - 2\tilde{x}_x \tilde{y}_x}{2\tilde{x}_x \tilde{y}_x} \right) + \left(\frac{\tilde{x}_x \tilde{y}_x + \tilde{x}_x \tilde{y}_y + \tilde{x}_y \tilde{y}_x + \tilde{x}_y \tilde{y}_y}{2\tilde{x}_x \tilde{y}_x} \right) = 0 \tag{25}$$

the result with the smaller absolute value $|\alpha_2|$ is selected.

3.3.2 Scale Correction

Once the shearing parameters have been obtained, the global scale can be chosen. To preserve as much information as possible, the number of pixels in image b and $\tilde{b}$ should be equal. The resolutions can be estimated from the length of the diagonal line through b and its projection in $\tilde{b}$. Using

$$\begin{gathered} \mathbf{a} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}}, \quad \mathbf{b} = \begin{bmatrix} W & H & 1 \end{bmatrix}^{\mathrm{T}} \\ \tilde{\mathbf{a}} = \begin{bmatrix} \alpha_2 & 0 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{H}_b^* \cdot \mathbf{a}, \quad \tilde{\mathbf{b}} = \begin{bmatrix} \alpha_2 & 0 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{H}_b^* \cdot \mathbf{b} \end{gathered} \tag{26}$$

one obtains the common scaling factor:

$$\alpha_3 = \frac{\left| \tilde{\mathbf{b}} - \tilde{\mathbf{a}} \right|}{\left| \mathbf{b} - \mathbf{a} \right|} \tag{27}$$

3.3.3 Mirroring Correction

To correct potential mirror effects, the order of the point correspondences are examined. First, the four corresponding triplets with the smallest and highest x- and y-value in image b are selected, to avoid flips of points due to perspective projection. If the order of the transformed values switches in one dimension for all three images, the rectified images have to be mirrored in that direction by setting m_x or m_y to -1.

3.3.4 Offset Estimation

The offsets s_1, s_2 and s_3 depend on the origin of the coordinate system. The estimate for s_1 is calculated from the rectification of the origins in image b and v. The parameter s_2 can be determined from the origin in image b and h. First, the homographies with shearing-, mirroring- and scaling correction are applied to the origins:

$$\mathbf{x}_b^* = \begin{bmatrix} x_{b1}^* \\ x_{b2}^* \\ x_{b3}^* \end{bmatrix} = \begin{bmatrix} m_x\alpha_3 & 0 & 0 \\ 0 & m_y\alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \alpha_2 & 0 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{H}_b^* \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$
$$\mathbf{x}_h^* = \begin{bmatrix} x_{h1}^* \\ x_{h2}^* \\ x_{h3}^* \end{bmatrix} = \begin{bmatrix} m_x\alpha_3 & 0 & 0 \\ 0 & m_y\alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1-\alpha_1 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{H}_h^* \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (28)$$
$$\mathbf{x}_v^* = \begin{bmatrix} x_{v1}^* \\ x_{v2}^* \\ x_{v3}^* \end{bmatrix} = \begin{bmatrix} m_x\alpha_3 & 0 & 0 \\ 0 & m_y\alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \alpha_2 & 0 & 0 \\ 1-\alpha_2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{H}_v^* \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Now, the horizontal and vertical offsets are defined by the negation of the minimal coordinates:

$$s_1 = -\min(x_{b1}^*, x_{v1}^*)$$
$$s_2 = -\min(x_{b2}^*, x_{h2}^*) \quad (29)$$

The remaining parameter s_3 is calculated from the origin of image h and v. Since the parameters s_1 and s_2 are already known, we can use

$$\mathbf{y}_h^* = \begin{bmatrix} 1 & 0 & s_1 \\ 0 & 1 & s_2 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{h1}^* \\ x_{h2}^* \\ x_{h3}^* \end{bmatrix}$$
$$\mathbf{y}_v^* = \begin{bmatrix} 1 & 0 & s_1 \\ 0 & 1 & s_2 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{v1}^* \\ x_{v2}^* \\ x_{v3}^* \end{bmatrix} \quad (30)$$

to determine the missing parameter

$$s_3 = -\min(y_{v2}^*, y_{h1}^*) . \quad (31)$$

3.3.5 Finding the Common Region

Finally, the computation of the common image regions minimizes the image sizes and speed up image matching. Therefore, we cut off regions, which are not visible in all three images. The required clipping lines are shown in Figure 4.

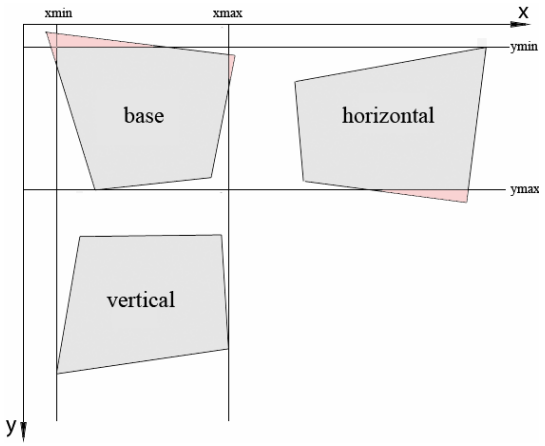

Figure 4. Definition of the common region

The vertical clipping line $\mathbf{x}_{min}$ can be derived from the corners of the images b and v using

$$\mathbf{a} = \mathbf{H}_b \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{b} = \mathbf{H}_b \cdot \begin{bmatrix} 0 \\ H \\ 1 \end{bmatrix}, \quad \mathbf{c} = \mathbf{H}_v \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{d} = \mathbf{H}_v \cdot \begin{bmatrix} 0 \\ H \\ 1 \end{bmatrix} \quad (32)$$
$$\mathbf{x}_{\min} = \begin{bmatrix} 1 & 0 & -\max(\min(\mathbf{a},\mathbf{b}), \min(\mathbf{c},\mathbf{d})) \end{bmatrix}^{\mathrm{T}}$$

The line $\mathbf{x}_{max}$ is given by

$$\mathbf{a} = \mathbf{H}_b \cdot \begin{bmatrix} W \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{b} = \mathbf{H}_b \cdot \begin{bmatrix} W \\ H \\ 1 \end{bmatrix}, \quad \mathbf{c} = \mathbf{H}_v \cdot \begin{bmatrix} W \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{d} = \mathbf{H}_v \cdot \begin{bmatrix} W \\ H \\ 1 \end{bmatrix} \quad (33)$$
$$\mathbf{x}_{\max} = \begin{bmatrix} 1 & 0 & -\min(\max(\mathbf{a},\mathbf{b}), \max(\mathbf{c},\mathbf{d})) \end{bmatrix}^{\mathrm{T}}$$

The horizontal clipping lines $\mathbf{y}_{min}$ and $\mathbf{y}_{max}$ are determined from the corners of the images b and h in an analogous manner.

3.4 Applying the Rectifying Transformation

To achieve optimal accuracy, the images b, h and v are resampled only once, using equation (16), by the indirect method. For every integer position in the rectified image, the non-integer position in the original image is determined using the inverse projective transformation $\mathbf{H}^{-1}$. The sub-pixel intensities are then computed by bicubic interpolation, which determines the intensity value from 16 pixels in a 4×4 neighborhood (see Figure 5). The computation

$$f(i+dx, j+dy) = \sum_{m=-1}^{2} \sum_{n=-1}^{2} f(i+m, j+n) \cdot r(m-dx) \cdot r(dy-n) \quad (34)$$

using the cubic weighting function

$$r(k) = \frac{1}{6}\left[p(k+2)^3 - 4p(k+1)^3 + 6p(k)^3 - 4p(k-1)^3 \right]$$
$$\text{with} \quad p(k) = \begin{cases} k & \text{for } k > 0 \\ 0 & \text{otherwise} \end{cases} \quad (35)$$

defines the interpolated intensity value and must be computed separately for each color channel.

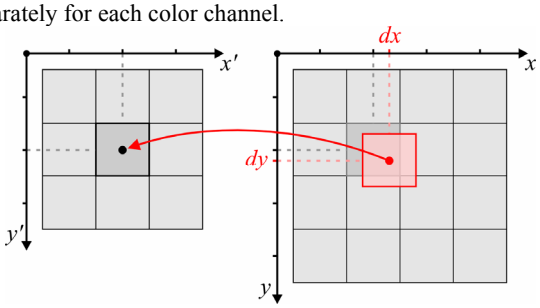

Figure 5. Resampling using the indirect transformation

4. EXPERIMENTAL RESULTS

The proposed method is verified by rectifying some real images of a statue, which were acquired by hand with an uncalibrated digital camera (see Figure 6). The original image size is 1024x768 pixels. For each image, 16 corresponding points were measured and the trifocal tensor was computed. The results are illustrated in Figure 7. To verify the correction of mirroring effects and the image alignment, several image triples were permutated and passed to the algorithm. The images were always positioned reasonably and no global mirror artefacts were observed.

The mean error and variance of the epipolar distance in the rectified images were computed for each image pair separately and for all three images. The results are shown in Table 1.

Distance	b↔h	b↔v	h↔v	Total
Mean error	0.378	0.285	0.573	0.398
Variance	0.374	0.163	0.688	0.393

Table 1: Epipolar distance of the rectified images [in pixels]

Figure 6. Image triplet before rectification

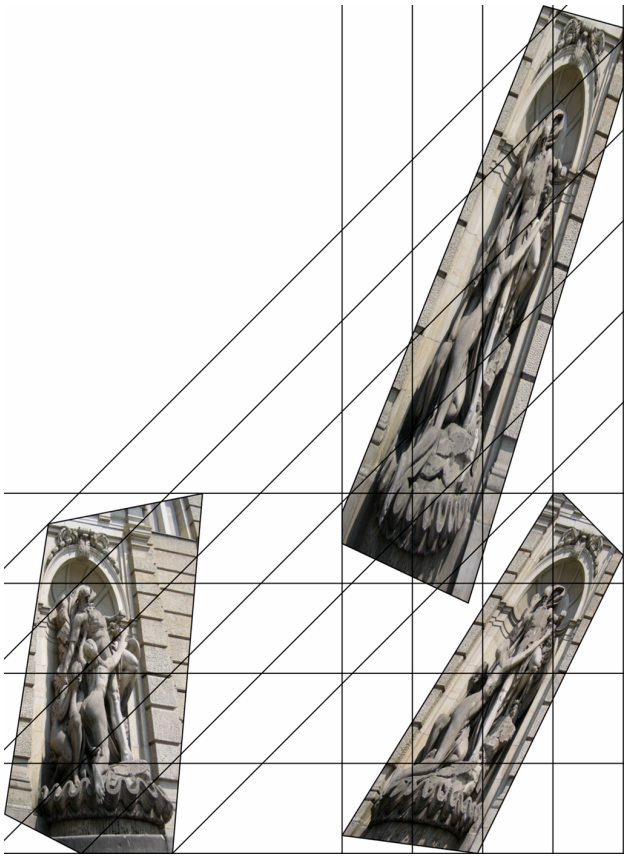

Figure 7. Image triplet after rectification

5. CONCLUSIONS

In this paper, a linear method for trinocular rectification of uncalibrated images, in closed form with 6 degrees of freedom, was proposed. In a post processing stage, proper geometric constraints are selected to minimize the projective distortion.
The proposed mirror correction eases the interpretation of the rectified images and makes it possible to apply this approach to various camera setups. Furthermore, the automated image alignment allows more convenient image acquisition, because the images can be shot without minding the relative image order. Finally, the computation of the common image regions minimizes the image sizes and speeds up image matching. However, the quality of the rectification depends on the robust estimation of the fundamental matrix. Therefore, the correspondence sets should be carefully chosen and well distributed over the scene setup.

6. ACKNOWLEDGEMENTS

This work was partially supported by grants from the German Research Foundation DFG.

7. REFERENCES

An, L., Jia, Y., Wang, J., Zhang, X. and Li, M., 2004: An efficient rectification method for trinocular stereovision, *Proc. Int. Conf. on Pattern Recognition*, vol. 4, Cambridge, pp. 56-59.

Ayache, N. and Hansen, C., 1988: Rectification of Images for Binocular and Trinocular Stereovision, *Proc. 9th International Conference on Pattern Recognition*, Rome, pp. 11-16.

Hartley, R.I., 1999: Theory and practice of projective rectification, *Int. Journal of Computer Vision*, vol. 35, no. 2, pp. 115-127.

Hartley, R.I. and Zisserman, A., 2000: *Multiple view geometry in computer vision*, Cambridge University Press, 607 p.

Heinrichs, M. and Rodehorst, V., 2006: Technical Report on Trinocular Rectification, Berlin University of Technology, http://www.cv.tu-berlin.de/publications/pdf/RectifyTR06.pdf

Loop, C. and Zhang, Z., 1999: Computing rectifying homographies for stereo vision, *Proc. Computer Vision and Pattern Recognition*, vol. 1, Fort Collins, pp. 125–131.

Matoušek, M., Šára, R. and Hlaváč, V., 2004: Data-optimal rectification for fast and accurate stereovision, *Proc. 3rd Int. Conf. on Image and Graphics*, Hong Kong, pp. 212-215.

Oram, D.T., 2001: Rectification for any epipolar geometry, *Proc. British Machine Vision Conf.*, London, pp. 653-662.

Rodehorst, V. and Hellwich, O., 2006: Genetic Algorithm SAmple Consensus (GASAC) - A Parallel Strategy for Robust Parameter Estimation, Int. Workshop "25 Years of RANSAC" in conjunction with CVPR '06, IEEE Computer Society, New York, 8 p.

Pollefeys, M., Koch, R. and Van Gool, L., 1999: A simple and efficient rectification method for general motion, *Proc. 7th Int. Conf. on Computer Vision*, vol. 1, Copenhagen, pp. 496–501.

Roy, S., Meunier, J. and Cox, I., 1997: Cylindrical rectification to minimize epipolar distortion, *Proc. Computer Vision and Pattern Recognition*, San Juan, pp. 393–399.

Sun, C., 2003: Uncalibrated Three-View Image Rectification, *Image and Vision Computing*, vol. 21, no. 3, pp. 259-269.

Torr, P.H.S. and Zisserman, A., 1997: Robust parameterization and computation of the trifocal tensor, *Image and Vision Computing*, vol. 15, no. 8, pp. 591-605

Zhang, H. and Šára, R. 2002: A linear method for trinocular rectification, Research Report No. 9, Center for Machine Perception, Czech Technical University Prague, 12 p.

Zhang, H., Čech, J., Šára, R., Wu, F. and Hu, Z., 2003: A Linear Method for Trinocular Rectification, *Proc. British Machine Vision Conference*, vol. 1, London, pp. 281-290.