A Linear Trinocular Rectification Method for Accurate Stereoscopic Matching

Huaifeng Zhang[†], Jan Čech[*], Radim Šára[*], Fuchao Wu[†], Zhanyi Hu[†]

[†]National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science
{hfzhang,fcwu,huzy}@nlpr.ia.ac.cn, http://www.nlpr.ia.ac.cn

[*]Center for Machine Perception, Faculty of Electrical Engineering, Czech Technical University
{cechj,sara}@cmp.felk.cvut.cz, http://cmp.felk.cvut.cz

Abstract

In this paper we propose and study a simple trinocular rectification method in which stratification to projective and affine components gives the rectifying homographies in a closed form. The class of trinocular rectifications which has 6 DOF is parametrized by an independent set of parameters with a geometric meaning. This offers the possibility to minimize rectification distortion in a natural way. It is shown experimentally on real data that our algorithm performs the rectification task correctly. As shown on ground-truth data using Confidently Stable Matching, trinocular matching significantly improves disparity map density and mismatch error, both depending on texture strength. Matching results on real complex scenes are reported.

1 Introduction

Rectification is often applied in two-view stereo matching. Given a pair of images, rectification determines a homography for each image mapping pairs of corresponding epipolar lines onto rectified image rows.

Rectification significantly simplifies area-based matching because (1) correlation statistics can be computed using efficient algorithms that use each image pixel only a few times, (2) non-local matching constraints (like ordering) are easier to represent, (3) sub-pixel and/or adaptive window matching is simplified because first-order approximation to projective distortion becomes a 2-parameter *affine* mapping [10], (4) 3D reconstruction becomes a linear problem and error propagation becomes easy [10].

The general binocular rectification problem is discussed by Hartley in [5]. In his paper, quasi-affine transformation is proposed to avoid rectified images being split. Pollefeys et al. [8] presented a rectification method from the geometric point of view. Their method can still work well when the epipoles are in the images. Loop and Zhang [6] presented a stratified method which deals with the rectification through a decomposition of the transformation into three parts. Gluckman and Nayar's [3] method minimizes the projective distortion by preserving the unit value of the Jacobian determinant. Roy et al. [9] proposed a rectification method that maps the images onto a cylinder. Their method guarantees the rectified images remain bounded for all possible camera motions. For a review of other binocular rectification methods, see [14].

It has long been recognized that multi-view matching can considerably improve the quality of the resulting disparity map [13, 2, 12]. This observation will be confirmed in this paper. Rectification remains of interest here: In multi-view matching, speeding up the computing time is even more important. Ayache and Hansen [1] gave one of the first methods for trinocular rectification. The result is dependent on the choice of the world coordinate system. In Sun's paper [14], several methods of trinocular rectification are discussed. Their rectification conditions are weaker than those we propose here.

We discuss trinocular rectification in which the rectification homographies can be computed in a closed form. We show that the rectification homographies have 6 degrees of freedom (DOF) and can be parametrized by independent parameters with a geometric meaning. This offers a possibility to minimize rectification distortion in a natural way.

2 Trinocular Rectification

We pose three constraints on trinocular rectification. In the subsequent text, we denote the rectified images as $\bar{b}$ (bottom, the bar means rectified image), $\bar{r}$ (right) and $\bar{t}$ (top), respectively. The three constrains can be described as follows (follow Fig. 1). Images $\bar{b}$, $\bar{r}$ and $\bar{t}$ are rectified if:

1. All the epipolar lines of images $\bar{b}$ and $\bar{r}$ are horizontal and the corresponding points have the same u-coordinate.
2. All the epipolar lines of images $\bar{b}$ and $\bar{t}$ are vertical and the corresponding points have the same v-coordinate.
3. For any triple of corresponding points $(\bar{x}_b, \bar{x}_r, \bar{x}_t)$, the disparity between $\bar{x}_b, \bar{x}_r$ and $\bar{x}_b, \bar{x}_t$ is equal. [1]

The rectification constraints can be represented by fundamental matrices. For convenience, we give the canonical fundamental matrices of the rectified images.

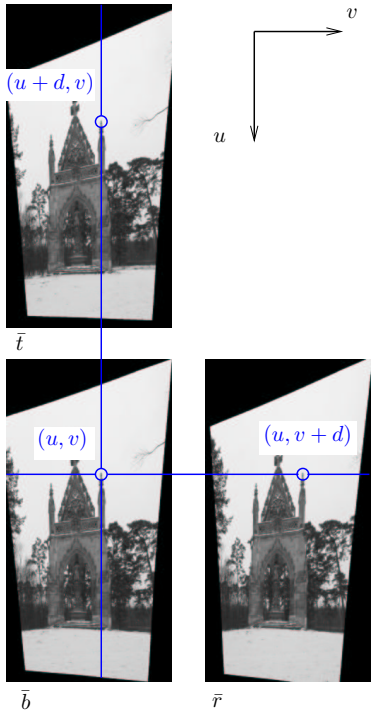


Figure. 1: A rectified image triple.

It is not difficult to verify that the constraints of trinocular rectification hold if the fundamental matrices between the rectified image pairs $\bar{b}\bar{r}$, $\bar{b}\bar{t}$, and $\bar{r}\bar{t}$ are

$$\bar{F}_{br} \simeq \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad \bar{F}_{bt} \simeq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \bar{F}_{rt} \simeq \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}, \tag{1}$$

where $\simeq$ means 'equal up to a non-zero scale.'

[1] As a result of this requirement different baselines result in image shrinkage in the direction of the longer baseline (as observed in Fig. 1). This makes the depth sensitivity to unit disparity error equal in both horizontal and vertical pair, which is desired. Algorithmic advantages are not of primary importance here.

In trinocular rectification, the final goal is to determine three rectification homographies that transform the images so that the rectification conditions hold with minimal image distortion. In other words one needs rectification homographies H_b, H_r, H_t that make the fundamental matrices have the form (1) and can be parameterized by a set of independent free parameters over which it is possible to minimize the distortion. Let u^T, v^T, w^T be the row vectors of H. The three homographies can be written as follows

$$H_b = \begin{bmatrix} u_b^T \\ v_b^T \\ w_b^T \end{bmatrix}, \quad H_r = \begin{bmatrix} u_r^T \\ v_r^T \\ w_r^T \end{bmatrix}, \quad H_t = \begin{bmatrix} u_t^T \\ v_t^T \\ w_t^T \end{bmatrix}. \tag{2}$$

Let x_b, x_r, x_t be a triple of corresponding points before rectification and $\bar{x}_b$, $\bar{x}_r$, $\bar{x}_t$ be the triple after rectification. Let the fundamental matrices before and after rectification be F_{br}, F_{bt}, F_{rt} and $\bar{F}_{br}$, $\bar{F}_{bt}$, $\bar{F}_{rt}$, respectively. We have

$$\bar{x}_b \simeq H_b \cdot x_b\,, \quad \bar{x}_r \simeq H_r \cdot x_r\,, \quad \bar{x}_t \simeq H_t \cdot x_t\,.$$

Then $\bar{x}_r^T \bar{F}_{br} \bar{x}_b = 0$ implies $x_r^T H_r^T \bar{F}_{br} H_b x_b = 0$ because $x_r^T F_{br} x_b = 0$. Similarly for the other two pairs. We therefore have

$$H_r^T \bar{F}_{br} H_b = \lambda_1 \cdot F_{br}\,, \quad H_t^T \bar{F}_{bt} H_b = \lambda_2 \cdot F_{bt}\,, \quad H_t^T \bar{F}_{rt} H_r = \lambda_3 \cdot F_{rt} \tag{3}$$

for some non-zero λ_1, λ_2, λ_3. Let us examine the degrees of freedom in this set of constraints. They will be the DOF of trinocular rectification. Obviously, the number of the independent elements in F_{br}, F_{bt} and F_{rt} is the number of the constraints on the unknown homographies. In three 3×3 fundamental matrices, there are altogether 27 elements. We know that fundamental matrices are given up to scale, the determinants $|F_{br}|$, $|F_{bt}|$, $|F_{rt}|$ all vanish, and there are three constrains among F_{br}, F_{bt} and F_{rt} [4, p. 378]. So there are altogether $27 - 3 - 3 - 3 = 18$ independent constraints represented by the three fundamental matrices.

Any 3×3 homography (up to scale) has 8 independent elements. So, for 3 homographies, there are $3 \times 8 = 24$ independent unknown parameters to be determined. Therefore the equation set (3) is an under-constrained problem with $24 - 18 = 6$ degrees of freedom.

3 A Linear Method for Trinocular Rectification

By combining (3), (1) and (2), we obtain

$$\begin{aligned} &-w_r \cdot u_b^T + u_r \cdot w_b^T = \lambda_1 \cdot F_{br}\,, \\ &w_t \cdot v_b^T - v_t \cdot w_b^T = \lambda_2 \cdot F_{bt}\,, \\ &-w_t (u_r + v_r)^T + (u_t + v_t) \cdot w_r^T = \lambda_3 \cdot F_{rt}\,. \end{aligned} \tag{4}$$

If w_b, w_r, w_t are known, the (4) becomes a *linear* set of equations for the remaining elements of the homographies (2).

3.1 Mapping the Epipoles to Infinity

For rectification, one of the necessary conditions is mapping all the epipoles to infinity. Let e_{br} and e_{rb} be the epipoles of the pair br, which lie in image planes b and r, respectively. Similarly, let e_{bt}, e_{tb} be the epipoles of the pair bt, and e_{rt}, e_{tr} be the epipoles of

the pair rt. We have

$$H_b \cdot e_{br} = \begin{bmatrix} u_b^T e_{br} \\ v_b^T e_{br} \\ w_b^T e_{br} \end{bmatrix} = \begin{bmatrix} \times \\ \times \\ 0 \end{bmatrix}, \qquad H_b \cdot e_{bt} = \begin{bmatrix} u_b^T e_{bt} \\ v_b^T e_{bt} \\ w_b^T e_{bt} \end{bmatrix} = \begin{bmatrix} \times \\ \times \\ 0 \end{bmatrix}.$$

From the above two equations, we get

$$w_b^T e_{br} = 0, \quad w_r^T e_{bt} = 0 \qquad \Rightarrow \qquad w_b \simeq e_{br} \times e_{bt}. \tag{5}$$

From the other two pairs we get

$$w_r \simeq e_{rb} \times e_{rt}, \quad w_t \simeq e_{tb} \times e_{tr}. \tag{6}$$

3.2 Primitive Rectification

Once we have the w_b, w_r, w_t, the equation set (3) can be solved as a linear one.

Proposition 1 *The following is a solution to* (3)*:*

$$\begin{aligned} H_b^* &= \begin{bmatrix} F_{33}^{br} w_{b1} - F_{31}^{br} & F_{33}^{br} w_{b2} - F_{32}^{br} & 0 \\ F_{31}^{bt} & F_{32}^{bt} & F_{33}^{bt} \\ w_{b1} & w_{b2} & 1 \end{bmatrix}, \\ H_r^* &= \begin{bmatrix} F_{13}^{br} & F_{23}^{br} & F_{33}^{br} \\ (F_{33}^{rt} + F_{33}^{br}) w_{r1} - F_{31}^{rt} - F_{13}^{br} & (F_{33}^{rt} + F_{33}^{br}) w_{r2} - F_{32}^{rt} - F_{23}^{br} & 0 \\ w_{r1} & w_{r2} & 1 \end{bmatrix}, \\ H_t^* &= \begin{bmatrix} (F_{33}^{br} - F_{33}^{bt}) w_{t1} + F_{13}^{rt} + F_{13}^{bt} & (F_{33}^{br} - F_{33}^{bt}) w_{t2} + F_{23}^{rt} + F_{23}^{bt} & F_{33}^{br} + F_{33}^{rt} \\ F_{33}^{bt} w_{t1} - F_{13}^{bt} & F_{33}^{bt} w_{t2} - F_{23}^{bt} & 0 \\ w_{t1} & w_{t2} & 1 \end{bmatrix}. \end{aligned} \tag{7}$$

We call this solution the primitive rectification homographies. From the last section, we know there are 6 DOF in (3). So we should select 6 parameters. Suppose u_{b3} is the 3rd element of u_b, similarly v_{t3} is the third element of v_t and v_{r3} of v_r. If we set $u_{b3} = v_{t3} = v_{r3} = 0$, $\lambda_1 = \lambda_2 = \lambda_3 = 1$, the expressions (7) follow.

Proposition 2 *Let* (H_b^*, H_r^*, H_t^*) *be the primitive solution to* (3)*. Then*

$$\begin{aligned} H_b &= \begin{bmatrix} 1 & 0 & s_1 \\ 0 & 1 & s_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_3 & 0 & 0 \\ 0 & \alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot H_b^*, \\ H_r &= \begin{bmatrix} 1 & 0 & s_1 \\ 0 & 1 & s_2 + s_3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_3 & 0 & 0 \\ 0 & \alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 & 0 & 0 \\ 1 - \alpha_1 & 1 & F_{33}^{br}(\alpha_1 - 1) \\ 0 & 0 & 1 \end{bmatrix} \cdot H_r^*, \\ H_t &= \begin{bmatrix} 1 & 0 & s_1 + s_3 \\ 0 & 1 & s_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_3 & 0 & 0 \\ 0 & \alpha_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 - \alpha_2 & F_{33}^{br}(\alpha_1 - 1) \\ 0 & \alpha_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot H_t^* \end{aligned} \tag{8}$$

are also solutions to (3)*, where* s_1, s_2, s_3 *and* α_1, α_2, α_3 *are free parameters.*

It is not difficult to obtain the form of (8) if we set $s_1 = u_{b3}$, $s_2 = v_{t3}$, $s_3 = v_{r3} - v_{t3}$ and $\alpha_1 = \lambda_1/\lambda_3$, $\alpha_2 = \lambda_2/\lambda_3$ and $\alpha_3 = \lambda_3$.

From (8), we can see that each of the six parameters has a geometric meaning:

1. the s_1, s_2 are the common u-shift and v-shift of the three images, respectively,
2. the s_3 is the shift in the v-direction of image r, which is equal to the shift in the u-direction of image t,
3. the α_1 is the scale in the u-direction of images b and r, which affects the shearing of image r, the v-shift of image r, and the u-shift of image t,
4. the α_2 is the scale in the v-direction of images b and t, which affects the shearing of image t,
5. the α_3 is the common scale in both directions of the three images.

The primitive rectification (H_b^*, H_r^*, H_t^*) can be applied to transform the three images to get a set of pre-rectified images denoted as $\hat{b}, \hat{r}$ and $\hat{t}$, respectively. The parameters s_1, s_2, s_3, α_1, α_2, α_3 can be selected by minimal distortion argument.

3.3 Distortion Correction

We need an optimal choice for the 6 DOF. Although $\hat{b}, \hat{r}$ and $\hat{t}$ are already rectified images, there may be a severe projective distortion present. We can correct it within the 6 DOF in two steps:

Mirroring correction Keeping other parameters fixed and changing the signs of α_1 and α_2, we may find a combination that preserves the order of the corner points of the three images.

Shearing correction Here, we try to select the proper values of α_1 and α_2 to make the shearing of images $\hat{r}$ and $\hat{t}$ minimal. Although α_1 also affects the translation of images r and t, we can deal with it afterwards. So in this step, we only consider the shearing effect of these two parameters.

Following the Loop and Zhang's paper [6] we minimize the shearing by preserving the perpendicularity of the two middle lines in the images. Let H and W be the height and width of the original images. Let $a = [0, W/2, 1]^T$, $b = [H, W/2, 1]^T$, $c = [H/2, 0, 1]^T$ and $d = [H/2, W, 1]^T$. Then, $\hat{a}_r = H_r^* \cdot a$, $\hat{b}_r = H_r^* \cdot b$,$\hat{c}_r = H_r^* \cdot c$ and $\hat{d}_r = H_r^* \cdot d$. Let $\hat{x} = \hat{a}_r - \hat{b}_r$, $\hat{y} = \hat{c}_r - \hat{d}_r$, and

$$S_r = \begin{bmatrix} \alpha_1 & 0 & 0 \\ \alpha_1 - 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The perpendicularity condition results in a quadratic equation in α_1

$$(S_r\hat{x})^T (S_r\hat{y}) = 0. \tag{9}$$

We select the solution with smaller $|\alpha_1|$. A similar condition gives us the solution for α_2. Next we select the α_3 to make the areas of images b and $\bar{b}$ equal. Finally, the remaining translation components s_1, s_2, s_3 are selected to obtain a proper bounding box for each rectified image.

4 Trinocular Confidently Stable Matching

In the experiments reported here we will be using a modified version of Confidently Stable Matching (CSM) [12]. The modification is as follows. Since the disparity in the $\overline{br}$ and $\overline{bt}$ pairs is the same, the matching costs

$$c(u,v,w) = \frac{1}{3}(c(u,v) + c(u,w) + c(u+w-v,v))$$

in the trinocular matching table are easily computed from three $n \times n$ windows in the bottom, right and top images, respectively, centered on positions (u,v), (u,w), and $(u+w-v,v)$, cf. Fig 1. The $c(\cdot,\cdot)$ values are the usual pairwise correlations. Modified normalized cross-correlation [7] was used in the experiments reported here.

In this approach all cameras play symmetric roles, none of them is a preferred camera, only the resulting disparity map is expressed in the bottom camera coordinates. The resulting algorithm will be denoted as 3CSM.

5 Experiments

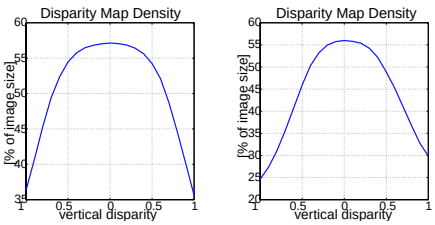
In this section we show that the proposed rectification is correct and that it is useful for obtaining accurate results in trinocular matching. We will first show results in an experiment with ground-truth and then results on two real scenes.

For the ground-truth experiment we collected images from a fully calibrated stereoscopic system with an L-shape camera arrangement. We used Pulnix TM-9701 cameras with Linos MeVis 25 mm high-quality lenses. The scene consisted of a tilted plane at approximately 1.2 m distance from the cameras. The horizontal and vertical camera baselines were 0.3 m and 0.2 m, respectively. A slide projector was used to project a random texture of optimal coarseness onto the test plane. Images were taken under three levels of projected texture contrast: strong (S), weak (W) and none (N), see Fig. 4. The raw image size was 484×760, after rectification it was 651×567 pixels. Ground-truth was obtained by robust fitting of a plane to the best-contrast trinocular disparity map. The CSM parameters were set to $\alpha = 10$, $\beta = 0.05$. Disparity range was $[20, 60]$. In all experiments we used 5×5 matching window.

The `StHubert` scene is an outdoor environment of great depth, many occlusions and fine details. The `Office` scene is a typical indoor environment with large regions of low texture contrast, slanted surfaces and repetitive structures. Both scenes were captured by a hand-held camera. Gray-scale images were used in matching. The three pairwise fundamental matrices were estimated by Zhang's method [15]. In both scenes the rectified image size was approximately 1200×700 and the expected disparity range in the interval of $[-100, 100]$.

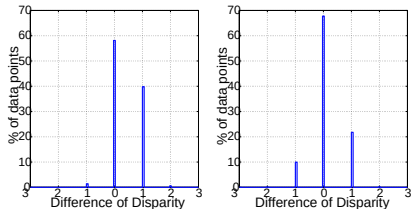
5.1 Verifying Trinocular Rectification Conditions

The goal of this experiment is to assess how far is the resulting rectification from the correct one. The following test does not guarantee that the images are rectified but it is simple and it at least checks the necessary conditions. We use the strong-texture data described above. In this experiment we kept the values $\alpha = 10$, $\beta = 0.05$.



(a) pair $\overline{br}$ (b) pair $\overline{bt}$

Figure 2: Matching density as a function of image shift.



(a) shift 0 (b) shift 0.3

Figure 3: Histograms of disparity difference $d_{br} - d_{bt}$.

First we consider the rectified images $\overline{b}$ and $\overline{r}$. The images are rectified, so every corresponding pair of pixels between $\overline{b}$ and $\overline{r}$ is on the same horizontal line. If the rectification is correct then by shifting one of the images along the vertical axis and repeating the dense matching, the matching quality must decrease. Since we are using the CSM algorithm [12] which guarantees a given mismatch probability at the expense of decreasing matching density and since we have an optimal surface texture, we can measure the matching quality as the density of the disparity map. Fig. 2 shows the density plots as functions of the shift for the $\overline{br}$ and the $\overline{bt}$ pairs (in the $\overline{bt}$ pair the corresponding shift is done in the horizontal direction). We can see that the matching quality peaks at zero shifts as expected.

The next check verifies that the vertical and horizontal disparities are equal for every triple of matched points. We compute the disparity map d_{br} of the horizontal pair $\overline{br}$ and the disparity map d_{bt} of the vertical pair $\overline{bt}$ and evaluate the difference $d_{br} - d_{bt}$. Fig. 3(a) shows the histogram of $d_{br} - d_{bt}$. From the histogram, we can see that most of the corresponding disparities are equal.

Note the significant fraction of disparities that differ by unity in Fig. 3(a). The observed differences suggest that the image discretization may play a negative role. To test this hypothesis we simultaneously shifted the right and top images by 0.3 pixels horizontally (vertically, respectively) and repeated the computation of the disparity difference. Note that this shift preserves the rectification condition. The improved result, shown in Fig. 3(b) confirms the large sensitivity of the trinocular matching algorithm to image discretization. Note that the cameras were calibrated very precisely and high-quality lenses without radial distortion were used. We conclude that subpixel image shift must be optimized for in trinocular rectification.

5.2 Improvement of Trinocular over Binocular Matching

As discussed in Sec. 5.1 trinocular rectification is more sensitive to discretization than binocular matching. The goal of this experiment is to show that despite this sensitivity matching results obtained from a triple of rectified images is consistently more accurate (in terms of mismatches) than the union of pairwise matching.

We run the binocular CSM algorithm on the $\overline{br}$, $\overline{bt}$ and $\overline{tr}$ pairs individually, the results are shown in Fig. 4 in the second to fourth columns. We then used stable union [11] for

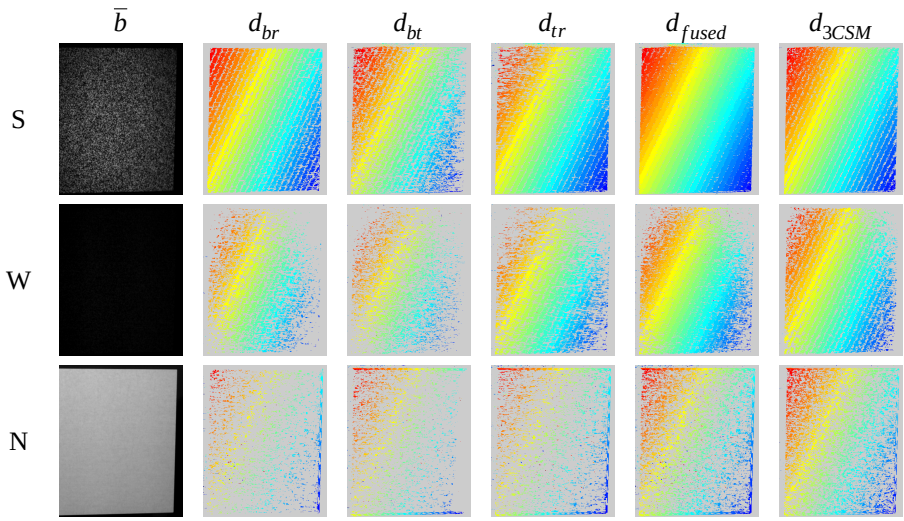


Figure 4: Disparity maps in the ground-truth experiment. Disparity is color-coded and gray marks unassigned disparities (holes).

	matching density [%]					**gross error rate** [%]				
	d_{br}	d_{bt}	d_{tr}	d_{fused}	d_{3CSM}	d_{br}	d_{bt}	d_{tr}	d_{fused}	d_{3CSM}
S	74	52	76	96	82	0.009	0.057	0.82	0.044	0
W	33	24	46	64	63	0.11	2	1.2	0.79	0.0042
N	10	10	12	26	33	4.7	3.9	7.1	5.4	0.58

Table 1: Evaluation results for strong (S), weak (W) and no-texture (N) images.

fusing the pairwise disparity maps as shown in the fifth column of Fig. 4. Finally the trinocular matching 3CSM was run to obtain the results in the last column. The rows are showing the results for three levels of texture contrast: strong (S), weak (W) and no-projected-texture (N).

To evaluate the matching performance, we define two types of error: (1) *matching density* as the percentage of image pixels with assigned disparity, (2) *gross error rate* as the percentage of matched pixels with disparity differing more than or equal to 2 pixels from the ground truth.

The results shown in Tab. 1 confirm the correctness of the trinocular matching and also show a significant improvement when trinocular matching is applied. Under weak texture trinocular matching is denser and much more accurate than any binocular matching or the fusion of binocular matchings. The stable union is denser under strong texture, but gross errors are often copied to the resulting disparity map. So, the quality of trinocular matching is provably better than just a plain fusion of binocular matchings.

5.3 Real Scenes

The results in the `StHubert` (first row) and the `Office` scenes (second row) are shown in Fig. 5. We used $\alpha = 20$, $\beta = 0.02$. Note that in the `StHubert` scene the trinocular

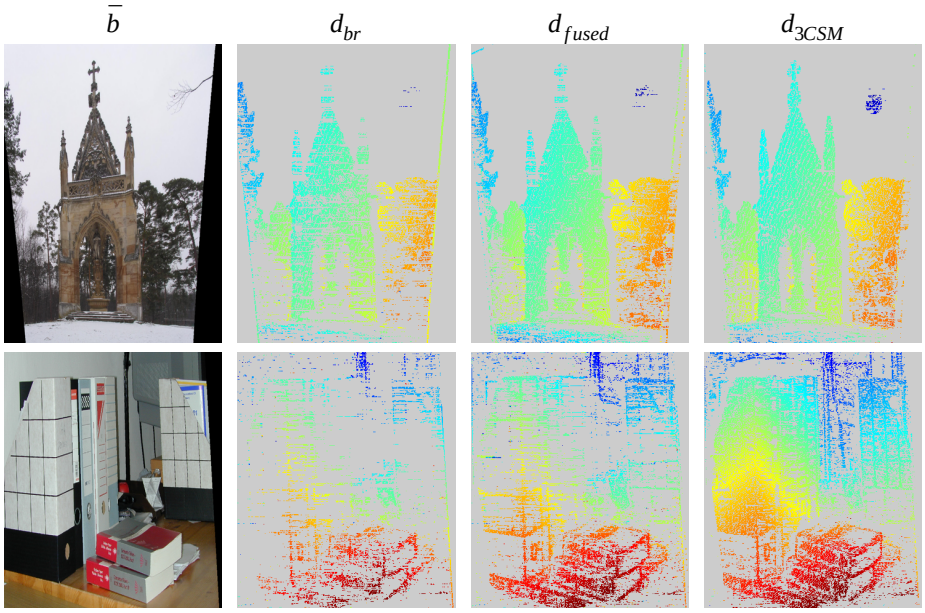


Figure 5: Matching results on real scenes.

algorithm picks up significantly more detail including the fine branch at the foreground (top right, in blue). There are noticeably fewer mismatches in the trinocular results than in the fusion of pairwise disparity maps. The trinocular map is denser with more crisp occlusion boundaries.

In the `Office` scene we see a great improvement in trinocular map density over the binocular case, especially in low-texture regions. Note fewer mismatches due to repetitive structures (e.g. the lines drawn on the paper folders).

6 Summary and Discussion

In this paper, a linear closed-form method for trinocular rectification was proposed. It was shown that trinocular rectification conditions leaves 6 DOF. The class of trinocular rectifications was parameterized by six independent parameters. By selecting their proper values, the projective distortion of trinocular rectification can easily be minimized in the stage of post-rectification.

The advantage of rectification is that the images are resampled *individually*, only *once*, with uniform interpolation neighborhood, and so *optimally* from signal-theoretic point of view. The rectified images thus have the optimal accuracy for matching.

Under general camera position, rectification is only possible for 2- or 3-camera stereo. As shown in this paper, the quality of dense trinocular matching is quite high and therefore a polynocular dense matching problem can be decomposed into independent 3-camera problems and the partial results then can be easily fused as in [10], for instance.

We assumed non-degenerate camera positions. When the optical centers of the three cameras fall on the same line the method will not work, since (5) and (6) all vanish. (We

can therefore check if the rectification is possible.)

In the neighborhood of epipoles the sensitivity of 3D reconstruction to correspondence error is unbounded and therefore the epipole area must be excluded from matching. Once this is done, the rest of the images can be split into n slices that are rectified independently.

In this method, the three images were indexed as b, r and t. Permutations of the indices will result in different rectifications. In this paper we did not deal with the problem of selecting the best indexing.

In the matching we considered only one type of occlusions. We assumed that a scene point is visible either in all three cameras or occluded. We did not distinguish all trinocular occlusion types. A correct occlusion handling is a topic for further research.

Acknowledgements The work was supported by the Czech Ministry of Education under projects MSMT Kontakt ME412 and MSM 210000012, by the National Science Foundation of China under project 60275009, and by the Grant Agency of the Czech Republic under project GACR 102/01/1371.

References

[1] N. Ayache and C. Hansen. Rectification of images for binocular and trinocular stereovison. In *ICPR*, pages 11–16, 1988.

[2] C. Buehler, J. Gortler, Steven, and L. McMillan. Minimal surfaces for stereo. In *ECCV*, pages 885–899, 2002.

[3] J. Gluckman and S. Nayar. Rectifying transformations that minimize resampling effects. In *CVPR*, 2001.

[4] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Kluwer Academic Publishers, 2000.

[5] R. I.Hartley. Theory and practice of projective rectification. *IJCV*, 35(2):115–127, 1999.

[6] C. Loop and Z. Zhang. Computing rectifying homographies for stereo vision. In *CVPR*, 1999.

[7] H. P. Moravec. Towards automatic visual obstacle avoidance. In *IJCAI*, page 584, 1977.

[8] M. Pollefeys, R. Koch, and L. Gool. A simple and efficient rectification method for general motion. In *ICCV*, 1999.

[9] S. Roy, J. Meunier, and I. Cox. Cylindrical rectification to minimize epipolar distortion. In *CVPR*, pages 393–399, 1997.

[10] R. Šára. Accurate natural surface reconstruction from polynocular stereo. In *NATO Advanced Research Workshop Confluence of Computer Vision and Computer Graphics*, number 84 in NATO Science Series 3., pages 69–86. Kluwer Academic Publishers, 2000.

[11] R. Šára. Sigma-delta stable matching for computational stereopsis. Research Report CTU–CMP–2001–25, Center for Machine Perception, K333 FEE Czech Technical University, 2001.

[12] R. Šára. Finding the largest unambiguous component of stereo matching. In *ECCV*, 2002.

[13] C. V. Stewart and C. R. Dyer. The trinocular general support algorithm: A three-camera stereo algorithm for overcoming binocular matching errors. In *ICCV*, pages 134–138, 1988.

[14] C. Sun. Uncalibrated three-view image rectification. *IVC*, 21(3):259–269, 2003.

[15] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *AI*, 78(1–2):87–119, 1995.