Quasi-Euclidean Uncalibrated Epipolar Rectification *

Andrea Fusiello
Dipartimento di Informatica, University of Verona
Strada Le Grazie, 15 - 37134 Verona, IT
andrea.fusiello@univr.it

Luca Irsara
EURAC research
Viale Druso 1, 39100 Bolzano, IT
luca.irsara@eurac.edu

Abstract

This paper deals with the problem of epipolar rectification in the uncalibrated case. First the calibrated (Euclidean) case is recognized as the ideal one, then we observe that in that case images are transformed with a collineation induced by the plane at infinity, which has a specific structure. That structure is therefore imposed to the sought transformation while minimizing the rectification error. Experiments show that this method yields images that are remarkably close to the ones produced by Euclidean rectification.

1. Introduction

Epipolar rectification is an important stage in dense stereo matching, as almost any algorithm requires rectified images, i.e., images where epipolar lines are parallel and horizontal and corresponding points have the same vertical coordinates. In the case of calibrated cameras the *Euclidean* epipolar rectification is unique up to trivial transformations [2].

On the contrary, in the case of uncalibrated cameras, there are more degrees of freedom in choosing the rectifying transformation [4] and a few competing methods are present in the literature [10, 11]. Each aims at producing a "good" rectification by minimizing a measure of distortion, but none is clearly superior to the others, not to mention the fact that there is no agreement on what the distortion criterion should be. Above all, none of them achieves results comparable to the Euclidean epipolar rectification, which can be indisputably taken as the target result. This paper aims at achieving a good approximation of the Euclidean epipolar rectification, which we refer to as *quasi-Euclidean* epipolar rectification.

Geometrically, in the Euclidean frame, rectification is achieved by a suitable rotation of both image planes. The correspondent image transformation is the collineation induced by the plane at infinity. As a result, the plane at infinity is the locus of zero-disparity in the rectified stereo pair. This is signified by saying that Euclidean rectification is done *with respect to* the plane at infinity. In the uncalibrated case the reference plane is generic, as any plane can play the role of the infinity plane in the projective space. Our quasi-Euclidean rectification can be seen as referred to a plane that approximates the plane at infinity.

1.1. Previous work

The first work on uncalibrated rectification (called "matched-epipolar projection") is [5], followed by [4], where the author tidies up the theory. He uses the condition that one of the two collineations should be close to a rigid transformation in the neighborhood of a selected point, while the remaining degrees of freedom are fixed by minimizing the distance between corresponding points (disparity). Along the same line, [12] also proposes a distortion criterion based on simple geometric heuristics.

Loop and Zhang [11] decompose each collineation into similarity, shearing and projective factors and attempt to make the projective component "as affine as possible". Isgrò and Trucco [10] build upon [4] and propose a method that avoids computation of the fundamental matrix, using the same distortion criterion as in [4]. The practice has shown that the rectification produced by these methods is not always satisfactory, if compared to results obtained in the calibrated case. Wu and Yu [9] argue that minimizing the disparity might be the cause of the problem, and propose a technique which is similar to [10] but uses a different distortion criterion derived from [11]. Rectification ends up in a non-linear minimization with six degrees of freedom. In [3] the transformation that best preserve the sampling of the original images is selected, by penalizing minification and magnification effects.

A different approach is followed in [1]: they design the collineations so as to minimize the relative distortion between the rectified images (instead of the distortion of each rectified image with respect to the original one), and the remaining degrees of freedom are fixed by choosing the refer-

*This research was partially supported by PRIN 2006 project 3-SHIRT.

ence plane in the scene, that will have zero disparity in the rectified pair. This choice, which affects sensibly the quality of rectification, is left to the user. The merit of this approach, however, is to make the role of the reference plane explicit.

2. Background

In this section we shall briefly recapitulate the theory of calibrated (or Euclidean) epipolar rectification; the reader is referred to [2] for more details.

Given two camera matrices P_{or} and $P_{o\ell}$, the idea behind rectification is to define two new *virtual* cameras P_{nr} and $P_{n\ell}$ obtained by rotating the actual ones around their optical centres until focal planes become coplanar. The rectification method describes how to compute the new cameras. Then (we concentrate on the right camera, but the same reasoning applies to the left one), the rectifying transformation that is to be applied to images is given by the 3×3 matrix:

$$H_r = P_{nr_{1:3}} P_{or_{1:3}}^{-1} \tag{1}$$

where the subscript denotes a range of columns.

It is easy to see that H_r is the collineation induced by the plane at infinity between the old and the new cameras, hence it can be written as:

$$H_r = K_{nr} R_r K_{or}^{-1} \tag{2}$$

where K_{or} and K_{nr} are the intrinsic parameters of the old and new camera respectively, and R_r is the rotation that is applied to the old camera in order to rectify it.

The rectified images are as if they were taken by two cameras related by a translation along the baseline. Hence, the zero-disparity plane (i.e., the reference plane) is at infinity.

3. Method

We shall henceforth concentrate on the uncalibrated case. We assume that intrinsic parameters are unknown and that a number of corresponding tie-points $\mathbf{m}_\ell^j \leftrightarrow \mathbf{m}_r^j$ are available. The method proceeds along the same line of [10]: it seek the collineations that make the tie-points satisfy the epipolar geometry of a rectified image pair.

The fundamental matrix of a rectified pair has a very specific form, namely it is the skew-symmetric matrix associated with the cross-product by the vector $\mathbf{u}_1 = (1, 0, 0)$:

$$[\mathbf{u}_1]_\times = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \tag{3}$$

Let H_r and H_ℓ be the unknown rectifying collineations. When they are applied to the corresponding tie-points $\mathbf{m}_\ell^j, \mathbf{m}_r^j$ respectively, the transformed points must satisfy the epipolar geometry of a rectified pair, namely:

$$(H_r \mathbf{m}_r^j)^\mathsf{T} [\mathbf{u}_1]_\times (H_\ell \mathbf{m}_\ell^j) = 0, \tag{4}$$

The left-hand side of Equation (4) is an algebraic error, i.e., it has no geometrical meaning, so we used instead the Sampson error [6], that is a first order approximation of the geometric reprojection error. The matrix $F = H_r^\mathsf{T} [\mathbf{u}_1]_\times H_l$ can be considered as the fundamental matrix between the original images, therefore, in our case, the Sampson error for the j-th correspondence is defined as:

$$E_S^j = \frac{({\mathbf{m}_r^j}^\mathsf{T} F \mathbf{m}_\ell^j)^2}{||[\mathbf{u}_3]_\times F \mathbf{m}_\ell^j||^2 + ||{\mathbf{m}_r^j}^\mathsf{T} F [\mathbf{u}_3]_\times||^2}$$

where $\mathbf{u}_3 = (0, 0, 1)$

As this equation must hold for any j, one obtains a system of non-linear equations $\{E_S^j = 0\}$ in the unknown H_r and H_ℓ. A least-squares solution can be obtained with the Levenberg-Marquardt algorithm, but he way in which H_r and H_ℓ are parametrized is crucial, and characterizes our approach with respect to the previous ones. We force the rectifying collineations to have the same structure as in the calibrated (Euclidean) case, i.e., to be collineations induced by the plane at infinity, namely

$$H_r = K_{nr} R_r K_{or}^{-1} \quad H_\ell = K_{n\ell} R_\ell K_{o\ell}^{-1}. \tag{5}$$

The old intrinsic parameters $(K_{o\ell}, K_{or})$ and the rotation matrices (R_ℓ, R_r) are unknown, whereas the new intrinsic parameters $(K_{n\ell}, K_{nr})$ can be set arbitrarily, provided that vertical focal length and vertical coordinate of the principal point are the same. Indeed, it is easy to verify that the matrix $K_{nr}^\mathsf{T} [\mathbf{u}_1]_\times K_{n\ell}$ is equal (up to scale) to $[\mathbf{u}_1]_\times$, provided that the second and third row of K_{nr} and $K_{n\ell}$ are the same. Hence it is not necessary to include the matrices K_{nr} and $K_{n\ell}$ in the parametrization.

Each collineation depends in principle on five (intrinsic) plus three (rotation) unknown parameters. The rotation of one camera along its X-axis, however, can be eliminated. Consider the matrix

$$F = K_{or}^{-1} R_r^\mathsf{T} [\mathbf{u}_1]_\times R_\ell K_{o\ell}^{-1}. \tag{6}$$

Let R'_r and R'_ℓ be the same matrices as R_r and R_ℓ after pre-multiplying with an arbitrary (but the same for both) rotation matrix about the X-axis. It is easy to verify that $R_r^\mathsf{T} [\mathbf{u}_1]_\times R_\ell = R_r'^\mathsf{T} [\mathbf{u}_1]_\times R'_\ell$. Geometrically, this coincide with rotating a rectified pair around the baseline, which do not alter the rectification, but, in a real camera, it affects the portion of the scene that is imaged. Accordingly, we set to zero the rotation around the X-axis of the left camera.

We further reduce the number of parameters by making an educated guess on the old intrinsic parameters: no skew, principal point in the centre of the image, aspect ratio equal to one. The only remaining unknowns are the focal lengths of both cameras. Assuming that they are identical and equal to α, we get:

$$K_{or} = K_{o\ell} = \begin{bmatrix} \alpha & 0 & w/2 \\ 0 & \alpha & h/2 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

where w and h are width and height (in pixel) of the image.

In summary, the two collineations are parametrized by six unknowns: five angles and the focal length α. As suggested in [8], the focal length is expected to vary in the interval $[1/3(w+h), 3(w+h)]$, so we consider instead the variable $\alpha' = \log_3(\alpha/(w+h))$ which varies in $[-1, 1]$.

The minimization (using MATLAB "lsqnonlin") is carried out starting with all the unknown variables set to zero. The derivatives have been computed analytically with the MATLAB "Symbolic Math Toolbox".

Finally, the the rectifying collineations are computed with Eq. (5). The new intrinsic parameters (K_{nr} and $K_{n\ell}$) are set equal to the old ones: $K_{nr} = K_{n\ell} = K_{o\ell}$, modulo a shift of the principal point, that might be necessary to center the rectified images in the customary image coordinate frame. Horizontal translation has no effect on the rectification, whereas vertical translation must be the same for both images.

4. Results

We present first some results of our rectification algorithm applied to the SYNTIM images[1]. A few (10-20) corresponding tie-points were selected manually, then the algorithm automatically produced the rectified images. It always converged to a solution, in our experiments, without special initialization.

Calibration data (provided with the images) were used for computing the ground truth Euclidean rectification[2]. Unfortunately, a quantitative measure of distortion that captures the desired behaviour of rectification in every respect does not exist. Therefore, we are forced to leave the assessment of our quasi-Euclidean rectification to visual comparison with the ground truth: As the reader can appreciate in Fig. 1, our results are remarkably close to the Euclidean rectification.

In [10] the rectified images are in some cases very distorted and distinctly dissimilar from the Euclidean case. In the experiments shown in [9] the distortion is considerably reduced, but our results are still closer to the Euclidean rectification. The only one example reported in [11] does not allow for a meaningful comparison: Our quasi-Euclidean rectification, however, produces a very similar result (not reported here for space reasons).

Table 1 reports the rectification error, computed with

$$e_{\text{rec}} = \frac{1}{N}\sqrt{\sum_j E_S^j} \quad (8)$$

where N is the number of corresponding pairs. It can be interpreted as the root mean squared distance (in pixel) from each point to its (horizontal) epipolar line.

Table 1. Rectification errors.

	e_{rec} [pixel]	
Image pair	quasi-Euclid.	Euclid.
Aout	0.1063	0.3685
BalMouss	0.0753	0.7046
Rubik	0.1059	0.5524
Tot	0.0623	0.0946

These figures, beside confirming that the pairs are indeed rectified, reveal that the error achieved by the quasi-Euclidean rectification are consistently smaller than in the Euclidean rectification. This makes sense, because the former explicitly minimizes this error, whereas the latter derives the rectifying collineations directly from the camera matrices.

We noted experimentally that, considering separately the angle parameters and the focal length as the independent variables, the location of the minimum of the cost function is fairly insensitive to the value of the focal length (i.e., the correct angles can be obtained even when α is far away from the ground truth), whereas, in certain cases, small perturbations of the angles can drift the minimum away from the ground truth for the focal length (without affecting the the quality of the rectification, though). This seems to be related to the fact that computing the focal length from two uncalibrated views [7, 13] is ill conditioned when the cameras are verging.

5. Conclusion

We presented a new method for the epipolar rectification of uncalibrated stereo pairs which approximates the Euclidean (calibrated) case by enforcing the rectifying transformation to be a collineation induced by the plane at infinity. The method is based on the minimization of a cost function that has only six degrees of freedom and does not need any specific initialization. The results are close to the target Euclidean rectification and compares favorably with state-of-the-art uncalibrated methods, in terms of distortion applied to the rectified images.

Given the general utility of rectification, a MATLAB toolkit is available on the web[3].

[1] Available from www-rocq.inria.fr/~tarel/syntim/paires.html

[2] MATLAB code available from profs.sci.univr.it/~fusiello/demo/rect/

[3] http://profs.sci.univr.it/~fusiello/demo/rect

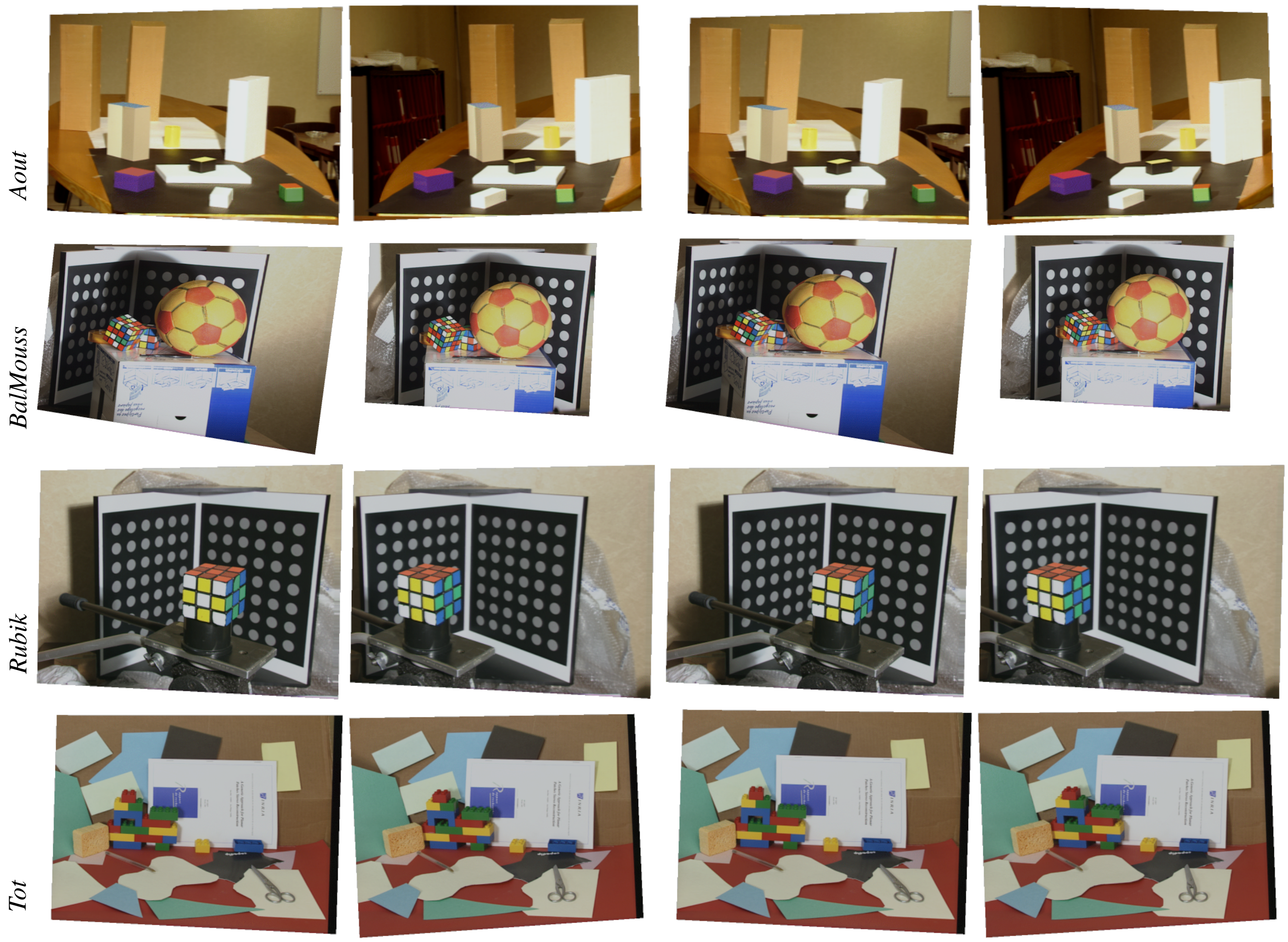

Figure 1. Euclidean (left) and quasi-Euclidean rectification (right) of the SYNTIM pairs.

References

[1] A. Al-Zahrani, S. S. Ipson, and J. G. B. Haigh. Applications of a direct algorithm for the rectification of uncalibrated images. *Information Sciences – Informatics and Computer Science*, 160(1-4):53–71, 2004.

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

[3] J. Gluckman and S. K. Nayar. Rectifying transformations that minimize resampling effects. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 01:111, 2001.

[4] R. Hartley. Theory and practice of projective rectification. *International Journal of Computer Vision*, 35(2):1–16, November 1999.

[5] R. Hartley and R. Gupta. Computing matched-epipolar projections. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 549–555, June 15-17 1993.

[6] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.

[7] R. I. Hartley. Estimation of relative camera position for uncalibrated cameras. In *Proceedings of the European Conference on Computer Vision*, pages 579–587, 1992.

[8] A. Heyden and M. Pollefeys. Multiple view geometry. In G. Medioni and S. B. Kang, editors, *Emerging Topics in Computer Vision*, pages 45 – 107. Prentice Hall, 2005.

[9] Y.-H. Y. Hsien-Huang P. Wu. Projective rectification with reduced geometric distortion for stereo vision and stereoscopic video. *Journal of Intelligent and Robotic Systems*, 42(1):Pages 71 – 94, Jan 2005.

[10] F. Isgrò and E. Trucco. Projective rectification without epipolar geometry. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages I:94–99, Fort Collins, CO, June 23-25 1999.

[11] C. Loop and Z. Zhang. Computing rectifying homographies for stereo vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages I:125–131, June 23-25 1999.

[12] L. Robert, M. Buffa, and M. Hebert. Weakly-calibrated stereo perception for rover navigation. In *ICCV*, pages 46–51, 1995.

[13] P. Sturm. On focal length calibration from two views. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume II, pages 145–150, 2001.