

A simple and efficient rectification method for general motion

Marc Pollefeys, Reinhard Koch¹ and Luc Van Gool
ESAT-PSI, K.U.Leuven
Kardinaal Mercierlaan 94
B-3001 Heverlee, Belgium
Marc.Pollefeys@esat.kuleuven.ac.be

Abstract

In this paper a new rectification method is proposed. The method is both simple and efficient and can deal with all possible camera motions. A minimal image size without any pixel loss is guaranteed. The only required information is the oriented fundamental matrix. The whole rectification process is carried out directly in the images. The idea consists of using a polar parametrization of the image around the epipole. The transfer between the images is obtained through the fundamental matrix. The proposed method has important advantages compared to the traditional rectification schemes. In some cases these approaches yield very large images or can not rectify at all. Even the recently proposed cylindrical rectification method can encounter problems in some cases. These problems are mainly due to the fact that the matching ambiguity is not reduced to half epipolar lines. Although this last method is more complex than the one proposed in this paper the resulting images are in general larger. The performance of the new approach is illustrated with some results on real image pairs.

1. Introduction

The stereo matching problem can be solved much more efficiently if images are rectified. This step consists of transforming the images so that the epipolar lines are aligned horizontally. In this case stereo matching algorithms can easily take advantage of the epipolar constraint and reduce the search space to one dimension (i.e. corresponding rows of the rectified images).

The traditional rectification scheme consists of transforming the image planes so that the corresponding space planes are coinciding [1]. There exist many variants of this traditional approach (e.g. [1, 5, 13]), it was even im-

plemented in hardware [2]. This approach fails when the epipoles are located in the images since this would have to results in infinitely large images. Even when this is not the case the image can still become very large (i.e. if the epipole is close to the image).

A recent approach by Roy, Meunier and Cox [15] avoids most of the problems of existing rectification approaches. This method proposes to rectify on a cylinder instead of a plane. It has two important features: it can deal with epipoles located in the images and it does not compress any part of the image. The method is however relatively complex since all operations are performed in 3-dimensional space. Every epipolar line is consecutively transformed, rotated and scaled. Although the method can work without calibration it still assumes oriented cameras. This is not guaranteed for a weakly calibrated image pair. The main disadvantage of this method is that it does not take advantage of the fact that the matching ambiguity in stereo matching is reduced to half epipolar lines. With the epipoles in the images this leads to an unnecessary computational overhead. For some cases this could even lead to the failure of stereo algorithms that enforce the ordering constraint. Although the size of the images is always finite, it is not minimal since the diameter of the cylinder is determined on the worst case pixel for the whole image.

Here we present a simple algorithm for rectification which can deal with all possible camera geometries. Only the oriented fundamental matrix is required. All transformations are done in the images. The image size is as small as can be achieved without compressing parts of the images. This is achieved by preserving the length of the epipolar lines and by determining the width independently for every half epipolar line.

For traditional stereo applications the limitations of standard rectification algorithms are not so important. The main component of camera displacement is parallel to the images for classical stereo setups. The limited vergence keeps the epipoles far from the images. New approaches in uncalibrated structure-from-motion [14] however make it possi-

¹Now at the Multimedia Information Processing, Inst. for Computer Science, Universität Kiel, Germany

ble to retrieve 3D models of scenes acquired with hand-held cameras. In this case forward motion can no longer be excluded. Especially when a street or a similar kind of scene is considered.

This paper is organized as follows. Section 2 introduces the concept of epipolar geometry on which the method is based. Section 3 describes the actual method. In Section 4 some experiments on real images show the results of our approach. The conclusions are given in Section 5.

2. Epipolar geometry

The epipolar geometry describes the relations that exist between two images. Every point in a plane that passes through both centers of projection will be projected in each image on the intersection of this plane with the corresponding image plane. Therefore these two intersection lines are said to be in epipolar correspondence.

This geometry can easily be recovered from the images [11] even when no calibration is available [6, 7]. In general 7 or more point matches are sufficient. Robust methods to obtain the fundamental matrix from images are described in [16, 18]. The epipolar geometry is described by the following equation:

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

where \mathbf{m} and \mathbf{m}' are homogenous representations of corresponding image points and \mathbf{F} is the fundamental matrix. This matrix has rank two, the right and left nullspace corresponds to the epipoles \mathbf{e} and \mathbf{e}' which are common to all epipolar lines. The epipolar line corresponding to a point \mathbf{m} is thus given by $\mathbf{l}' \sim \mathbf{F} \mathbf{m}$ with \sim meaning equality up to a non-zero scale factor (a strictly positive scale factor when oriented geometry is used, see further).

Epipolar line transfer The transfer of corresponding epipolar lines is described by the following equations:

$$\mathbf{l}' \sim \mathbf{H}^{-T} \mathbf{l} \text{ or } \mathbf{l} \sim \mathbf{H}^T \mathbf{l}' \quad (2)$$

with \mathbf{H} a homography for an arbitrary plane. As seen in [12] a valid homography can immediately be obtained from the fundamental matrix:

$$\mathbf{H} = [\mathbf{e}']_{\times} \mathbf{F} + \mathbf{e}' \mathbf{a}^T \quad (3)$$

with \mathbf{a} a random vector for which $\det \mathbf{H} \neq 0$ so that \mathbf{H} is invertible. If one disposes of camera projection matrices an alternative homography is easily obtained as:

$$\mathbf{H}^{-T} = (\mathbf{P}'^T)^{\dagger} \mathbf{P}^T \quad (4)$$

where \dagger indicates the Moore-Penrose pseudo inverse.

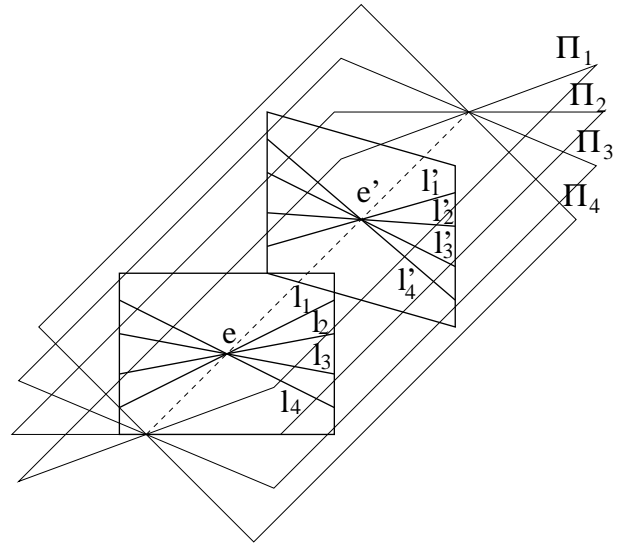


Figure 1. Epipolar geometry with the epipoles in the images. Note that the matching ambiguity is reduced to half epipolar lines.

Orienting epipolar lines The epipolar lines can be oriented such that the matching ambiguity is reduced to half epipolar lines instead of full epipolar lines. This is important when the epipole is in the image. This fact was ignored in the approach of Roy et al. [15].

Figure 1 illustrates this concept. Points located in the right halves of the epipolar planes will be projected on the right part of the image planes and depending on the orientation of the image in this plane this will correspond to the right or to the left part of the epipolar lines. These concepts are explained more in detail in the work of Laveau [10] on oriented projective geometry (see also [8]).

In practice this orientation can be obtained as follows. Besides the epipolar geometry one point match is needed (note that 7 or more matches were needed anyway to determine the epipolar geometry). An oriented epipolar line \mathbf{l} separates the image plane into a positive and a negative region:

$$f_1(\mathbf{m}) = \mathbf{l}^T \mathbf{m} \text{ with } \mathbf{m} = [x \ y \ 1]^T \quad (5)$$

Note that in this case the ambiguity on \mathbf{l} is restricted to a strictly positive scale factor. For a pair of matching points $(\mathbf{m}, \mathbf{m}')$ both $f_1(\mathbf{m})$ and $f_{1'}(\mathbf{m}')$ should have the same sign. Since \mathbf{l}' is obtained from \mathbf{l} through equation (2), this allows to determine the sign of \mathbf{H} . Once this sign has been determined the epipolar line transfer is oriented. We take the convention that the positive side of the epipolar line has the positive region of the image to its right. This is clarified in Figure 2.

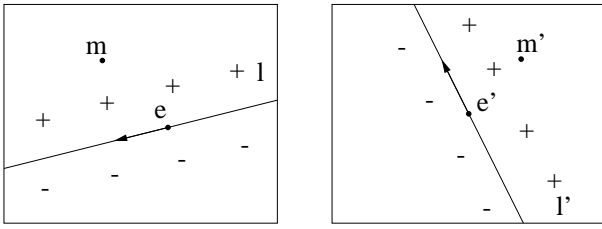


Figure 2. Orientation of the epipolar lines.

3. Generalized Rectification

The key idea of our new rectification method consists of reparameterizing the image with polar coordinates (around the epipoles). Since the ambiguity can be reduced to half epipolar lines only positive longitudinal coordinates have to be taken into account. The corresponding half epipolar lines are determined through equation (2) taking orientation into account.

The first step consists of determining the common region for both images. Then, starting from one of the extreme epipolar lines, the rectified image is build up line by line. If the epipole is in the image an arbitrary epipolar line can be chosen as starting point. In this case boundary effects can be avoided by adding an overlap of the size of the matching window of the stereo algorithm (i.e. use more than 360 degrees). The distance between consecutive epipolar lines is determined independently for every half epipolar lines so that no pixel compression occurs. This non-linear warping allows to obtain the minimal achievable image size without losing image information.

The different steps of this methods are described more in detail in the following paragraphs.

Determining the common region Before determining the common epipolar lines the extremal epipolar lines for a single image should be determined. These are the epipolar lines that touch the outer image corners. The different regions for the position of the epipole are given in Figure 3. The extremal epipolar lines always pass through corners of the image (e.g. if the epipole e is in region 1 the area between eb and ed). The extreme epipolar lines from the second image can be obtained through the same procedure. They should then be transferred to the first image. The common region is then easily determined as in Figure 4

Determining the distance between epipolar lines To avoid losing pixel information the area of every pixel should be at least preserved when transformed to the rectified image. The worst case pixel is always located on the image border opposite to the epipole. A simple procedure to compute this step is depicted in Figure 5. The same procedure

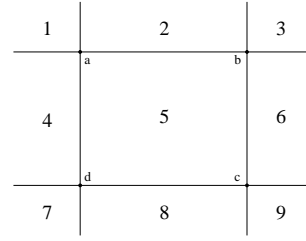


Figure 3. the extreme epipolar lines can easily be determined depending on the location of the epipole in one of the 9 regions. The image corners are given by a, b, c, d .

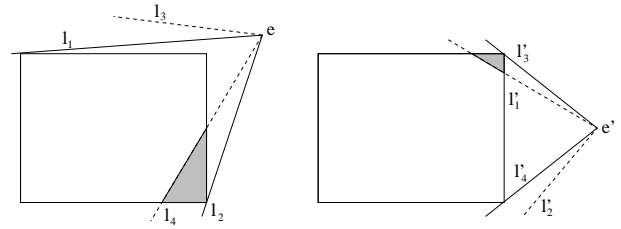


Figure 4. Determination of the common region. The extreme epipolar lines are used to determine the maximum angle.

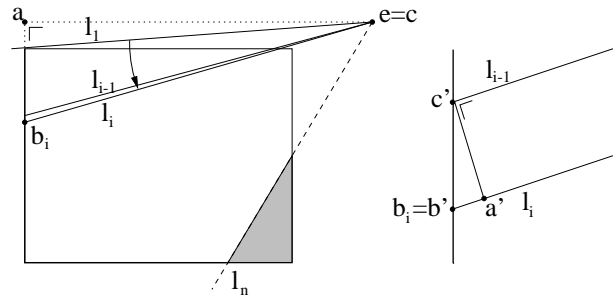


Figure 5. Determining the minimum distance between two consecutive epipolar lines. On the left a whole image is shown, on the right a magnification of the area around point b_i is given. To avoid pixel loss the distance $|a'c'|$ should be at least one pixel. This minimal distance is easily obtained by using the congruence of the triangles abc and $a'b'c'$. The new point b is easily obtained from the previous by moving $\frac{|bc|}{|ac|}$ pixels (down in this case).

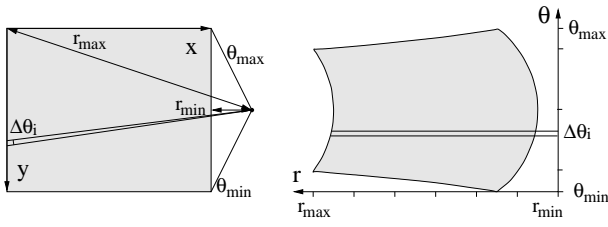


Figure 6. The image is transformed from (x,y) -space to (r,θ) -space. Note that the θ -axis is non-uniform so that every epipolar line has an optimal width (this width is determined over the two images).

can be carried out in the other image. In this case the obtained epipolar line should be transferred back to the first image. The minimum of both displacements is carried out.

Constructing the rectified image The rectified images are build up row by row. Each row corresponds to a certain angular sector. The length along the epipolar line is preserved. Figure 6 clarifies these concepts. The coordinates of every epipolar line are saved in a list for later reference (i.e. transformation back to original images). The distance of the first and the last pixels are remembered for every epipolar line. This information allows a simple inverse transformation through the constructed look-up table.

Note that an upper bound for the image size is easily obtained. The height is bound by the contour of the image $2 \times (W + H)$. The width is bound by the diagonal $\sqrt{W^2 + H^2}$. Note that the image size is uniquely determined with our procedure and that it is the minimum that can be achieved without pixel compression.

Transferring information back Information about a specific point in the original image can be obtained as follows. The information for the corresponding epipolar line can be looked up from the table. The distance to the epipole should be computed and subtracted from the distance for the first pixel of the image row. The image values can easily be interpolated for higher accuracy.

To warp a complete image back a more efficient procedure than a pixel-by-pixel warping can be designed. The image can be reconstructed radially (i.e. radar like). All the pixel between two epipolar lines can then be filled in at once from the information that is available for these epipolar lines. This avoids multiple look-ups in the table. More details on digital image warping can be found in [17].



Figure 7. Image pair from an Arenberg castle in Leuven scene.



Figure 8. Rectified image pair for both methods: standard homography based method (top), new method (bottom).

4. Experiments

As an example a rectified image pair from the Arenberg castle is shown for both the standard rectification and the new approach. Figure 7 shows the original image pair and Figure 8 shows the rectified image pair for both methods.

A second example shows that the methods works properly when the epipole is in the image. Figure 9 shows the two original images while Figure 10 shows the two rectified images. In this case the standard rectification procedure can not deliver rectified images.

A stereo matching algorithm was used on this image pair to compute the disparities. The raw and interpolate disparity maps can be seen in Figure 11. Figure 12 shows the depth map that was obtained and a few views of the result-

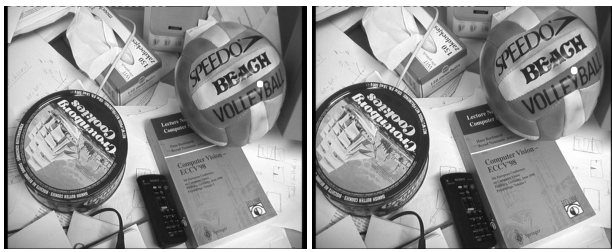


Figure 9. Image pair of the authors desk a few days before a deadline. The epipole is indicated by a white dot (top-right of 'Y' in 'VOLLEYBALL').

ing 3D model are shown in Figure 13. Note from these images that there is an important depth uncertainty around the epipole. In fact the epipole forms a singularity for the depth estimation. In the depth map of Figure 12 an artefact can be seen around the position of the epipole. The extend is much longer in one specific direction due to the matching ambiguity in this direction (see the original image or the middle-right part of the rectified image).

5. Conclusion

In this paper we have proposed a new rectification algorithm. Although the procedure is relatively simple it can deal with all possible camera geometries. In addition it guarantees a minimal image size. Advantage is taken of the fact that the matching ambiguity is reduced to half epipolar lines. The method was implemented and used in the context of depth estimation. The possibilities of this new approach were illustrated with some results obtained from real images.

Acknowledgments

We wish to acknowledge the financial support of the Belgian IUAP 24/02 'ImechS' project and of the EU ESPRIT ltr. 20.243 'IMPACT' project.

References

- [1] N. Ayache and C. Hansen, "Rectification of images for binocular and trinocular stereovision", *Proc. Intern. Conf. on Pattern Recognition*, pp. 11-16, 1988.
- [2] P. Courtney, N. Thacker and C. Brown, "A hardware architecture for image rectification and ground plane obstacle detection", *Proc. Intern. Conf. on Pattern Recognition*, pp. 23-26, 1992.
- [3] I. Cox, S. Hingorani and S. Rao, "A Maximum Likelihood Stereo Algorithm", *Computer Vision and Image Understanding*, Vol. 63, No. 3, May 1996.



Figure 10. Rectified pair of images of the desk. It can be verified visually that corresponding points are located on corresponding image rows. The right side of the images corresponds to the epipole.

- [4] L. Falkenhagen, "Hierarchical Block-Based Disparity Estimation Considering Neighbourhood Constraints", *Proc. International Workshop on SNHC and 3D Imaging*, Rhodes, Greece, 1997.
- [5] O. Faugeras, *Three-Dimensional Computer Vision: a Geometric Viewpoint*, MIT press, 1993.
- [6] O. Faugeras, "What can be seen in three dimensions with an uncalibrated stereo rig", *Computer Vision - ECCV'92*, Lecture Notes in Computer Science, Vol. 588, Springer-Verlag, pp. 563-578, 1992.
- [7] R. Hartley, "Estimation of relative camera positions for uncalibrated cameras", *Computer Vision - ECCV'92*, Lecture Notes in Computer Science, Vol. 588, Springer-Verlag, pp. 579-587, 1992.
- [8] R. Hartley, "Cheirality invariants", *Proc. D.A.R.P.A. Image Understanding Workshop*, pp. 743-753, 1993.
- [9] R. Koch, *Automatische Oberflächenmodellierung starrer dreidimensionaler Objekte aus stereoskopischen Rundum-Ansichten*, PhD thesis, University of Hannover, Germany, 1996 also published as Fortschritte-Berichte VDI, Reihe 10, Nr.499, VDI Verlag, 1997.
- [10] S. Laveau and O. Faugeras, "Oriented Projective Geometry for Computer Vision", in : B. Buxton and R. Cipolla (eds.),

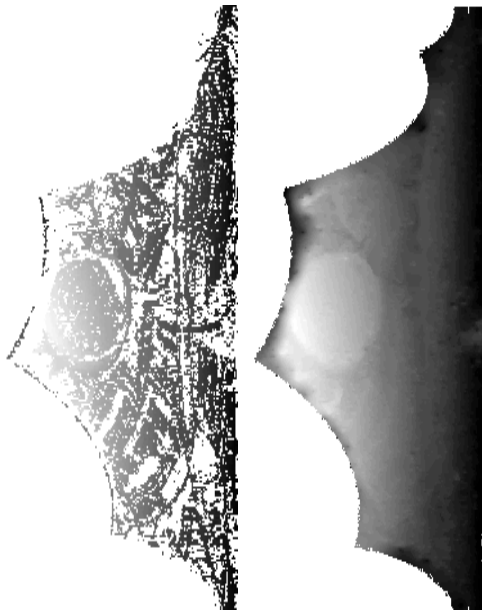


Figure 11. Raw and interpolated disparity estimates for the far image of the desk image pair.

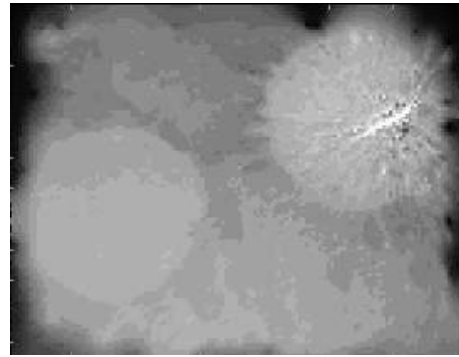


Figure 12. Depth map for the far image of the desk image pair.

Computer Vision - ECCV'96, Lecture Notes in Computer Science, Vol. 1064, Springer-Verlag, pp. 147-156, 1996.

- [11] H. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections", *Nature*, 293:133-135, 1981.
- [12] Q.-T. Luong and O. Faugeras, "The fundamental matrix: theory, algorithms, and stability analysis", *International Journal of Computer Vision*, 17(1):43-76, 1996.
- [13] D. Papadimitriou and T. Dennis, "Epipolar line estimation and rectification for stereo image pairs", *IEEE Trans. Image Processing*, 5(4):672-676, 1996.
- [14] M. Pollefeys, R. Koch, M. Vergauwen and L. Van Gool, "Metric 3D Surface Reconstruction from Uncalibrated Image Sequences", *Proc. SMILE Workshop (post-ECCV'98)*, Lecture Notes in Computer Science, Vol. 1506, pp.138-153, Springer-Verlag, 1998.
- [15] S. Roy, J. Meunier and I. Cox, "Cylindrical Rectification to Minimize Epipolar Distortion", *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp.393-399, 1997.
- [16] P. Torr, *Motion Segmentation and Outlier Detection*, PhD Thesis, Dept. of Engineering Science, University of Oxford, 1995.
- [17] G. Wolberg, *Digital Image Warping*, IEEE Computer Society Press Monograph, ISBN 0-8186-8944-7, 1990.
- [18] 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", *Artificial Intelligence Journal*, Vol.78, pp.87-119, October 1995.



Figure 13. Some views of the reconstruction obtained from the desk image pair. The inaccuracy at the top of the volleyball is due to the depth ambiguity at the epipole.