HERIOT-WATT
UNIVERSITY
EDINBURGH

DEPARTMENT of COMPUTING
and
ELECTRICAL ENGINEERING

Rectification with Unconstrained Stereo Geometry

A. Fusiello, E. Trucco and A. Verri

RESEARCH MEMORANDUM

Rectification with unconstrained stereo geometry

A. Fusiello, E. Trucco and A. Verri

Abstract

We present a linear rectification algorithm for general, unconstrained stereo rigs. The algorithm takes the two perspective projection matrices of the original cameras, and computes a pair of rectifying projection matrices. We report tests proving the correct behaviour of our method, as well as the negligible decrease of the accuracy of 3-D reconstruction if performed from the rectified images directly. To maximise reproducibility and usefulness, we give a working, 20-line MATLAB code, and a URL where the code can be found. Stereo reconstruction systems are very popular in vision research and applications, hence the usefulness of a general and easily accessible rectification algorithm.

1 Introduction and motivations

Given a pair of stereo images, *rectification* determines a transformation of each image plane such that pairs of conjugate epipolar lines become collinear and parallel to one of the image axes. The rectified images can be thought of as acquired by a new stereo rig, obtained by rotating the original cameras. The important advantage of rectification is that computing correspondences, a 2-D search problem in general, is reduced to a 1-D search problem, typically along the horizontal raster lines of the rectified images [9, 10, 7, 3, 16, 6].

This paper presents a novel algorithm rectifying a calibrated stereo rig of unconstrained geometry and mounting general cameras. The only input required is the pair of perspective projection matrices (PPM) of the two cameras (but not their individual intrinsic or extrinsic parameters), which can be estimated using one of the many existing calibration methods [5, 17, 2, 12]. The output is the pair of *rectifying* PPMs, which can be used to compute the rectified images. Reconstruction can also be performed directly from the rectified images and PPMs. Given the importance of rectification as a module for stereo systems, and the shortage of easily reproducible, easily accessible and clearly stated algorithms, we have made a "rectification kit" (code, examples data and instructions) available on line[1].

Some authors report rectification under restrictive assumptions; for instance, [11] assumes a very restrictive geometry (parallel vertical axes of the camera reference frames). Recently, [8, 13] have introduced an algorithm which performs rectification with general stereo geometry, given a *weakly calibrated* stereo rig (i.e. only the fundamental matrix is known).

This paper is organised as follows. Section 2 introduces our notations and summarises some necessary mathematics of perspective projections. Section 3 derive the algorithm for computing the rectifying PPMs and Section 3.1 expresses the rectifying image transformation in terms of PPMs. Section 4 gives the compact (20 lines), working MATLAB code for our algorithm, and indicates where to find our "rectification kit" on line. Section 5 reports tests on synthetic and real data. Section 6 is a brief discussion of our work.

[1] http://www.dimi.uniud.it/~fusiello.html

2 Camera model and epipolar geometry

This section briefly recalls the mathematical background on perspective projections necessary for our purposes. For more details see [4].

2.1 Camera model

A pinhole camera is modelled by its *optical centre* C and its *retinal plane* (or *image plane*) $\mathcal{R}$. A 3-D point W is projected into an image point M given by the intersection of the line containing C and W with $\mathcal{R}$ (see Fig. 1).

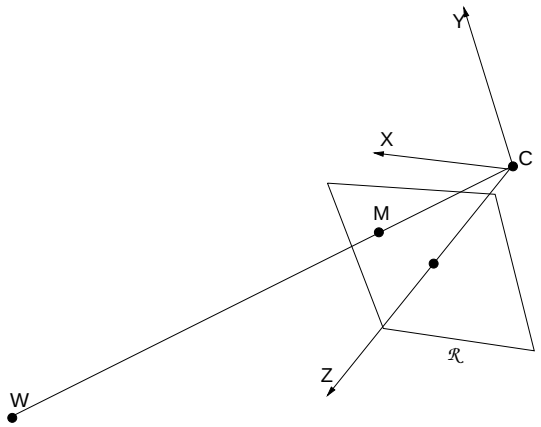

Figure 1: The pinhole camera model, with the *camera reference frame* (X,Y,Z) depicted. Z is also called the *optical axis*.

Let $\mathbf{w} = (x, y, z)$ the coordinates of W in the world reference frame (fixed arbitrarily) and $\mathbf{m}$ the coordinates of M in the image plane (pixels). In homogeneous (or projective) coordinates

$$\tilde{\mathbf{m}} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \qquad \tilde{\mathbf{w}} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{1}$$

the transformation from $\tilde{\mathbf{w}}$ to $\tilde{\mathbf{m}}$ is given by the matrix $\tilde{\mathbf{P}}$:

$$\lambda \tilde{\mathbf{m}} = \tilde{\mathbf{P}} \tilde{\mathbf{w}}, \tag{2}$$

where λ is an arbitrary scale factor.
The camera is therefore modelled by its *perspective projection matrix* (henceforth PPM) $\tilde{\mathbf{P}}$, which can be decomposed, using the QR factorisation, into the product

$$\tilde{\mathbf{P}} = \mathbf{A}[\mathbf{R} \mid \mathbf{t}]. \tag{3}$$

The matrix $\mathbf{A}$ depends on the *intrinsic parameters* only, and has the following form:

$$\mathbf{A} = \begin{bmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

where $\alpha_u = -fk_u$, $\alpha_v = -fk_v$ are the focal lengths in horizontal and vertical pixels, respectively (f is the focal length in millimeters, k_u and k_v are the effective number of pixels per millimiter along the u and v axes), (u_0, v_0) are the coordinates of the *principal point*, given by the intersection of the optical axis with the retinal plane, and γ is the *skew* factor.
The camera position and orientation (*extrinsic parameters*), are encoded by the 3×3 rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$, representing the rigid transformation that brings the camera reference frame onto the world reference frame.
Let us write the PPM as

$$\tilde{\mathbf{P}} = \left[\begin{array}{c|c} \mathbf{q}_1^\top & q_{14} \\ \mathbf{q}_2^\top & q_{24} \\ \mathbf{q}_3^\top & q_{34} \end{array}\right] = [\mathbf{Q}|\tilde{\mathbf{q}}]. \tag{5}$$

In cartesian coordinates, the projection (2) writes

$$\begin{cases} u = \dfrac{\mathbf{q}_1^\top \mathbf{w} + q_{14}}{\mathbf{q}_3^\top \mathbf{w} + q_{34}} \\ v = \dfrac{\mathbf{q}_2^\top \mathbf{w} + q_{24}}{\mathbf{q}_3^\top \mathbf{w} + q_{34}}. \end{cases} \tag{6}$$

The *focal plane* (the plane XY in Fig. 1) is parallel to the retinal plane and contains the optical centre. It is the locus of the points projected to infinity, hence its equation is $\mathbf{q}_3^\top \mathbf{w} + q_{34} = 0$. The two planes defined by $\mathbf{q}_1^\top \mathbf{w} + q_{14} = 0$ and $\mathbf{q}_2^\top \mathbf{w} + q_{24} = 0$ intersect the retinal plane in the vertical and horizontal axis of the retinal coordinates, respectively.
The optical centre C is the intersection of the these three planes, hence its coordinates $\mathbf{c}$ are the solution of

$$\tilde{\mathbf{P}} \begin{bmatrix} \mathbf{c} \\ 1 \end{bmatrix} = \mathbf{0} \tag{7}$$

then

$$\mathbf{c} = -\mathbf{Q}^{-1}\tilde{\mathbf{q}}. \tag{8}$$

From the latter a different way of writing $\tilde{\mathbf{P}}$ is obtained:

$$\tilde{\mathbf{P}} = [\mathbf{Q}| - \mathbf{Q}\mathbf{c}]. \tag{9}$$

The *optical ray* associated to an image point M is the line M C, i.e. the set of 3D points $\{\mathbf{w} : \tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{w}}\}$. The equation of this ray can be written in parametric form as

$$\mathbf{w} = \mathbf{c} + \lambda \mathbf{Q}^{-1}\tilde{\mathbf{m}} \quad \lambda \in \mathbb{R}. \tag{10}$$

2.2 Epipolar geometry

Let us consider a stereo rig composed by two pinhole cameras (Fig. 2). Let C_1 and C_2 be the optical centres of the left and right cameras respectively. A 3D point W is projected onto both image planes, to points M_1 and M_2, which constitute a conjugate pair. Given a point M_1 in the left image plane, its conjugate point in the right image is constrained to lie on a line called the *epipolar line* (of M_1). Since M_1 may be the projection of an arbitrary point on its optical ray, the epipolar

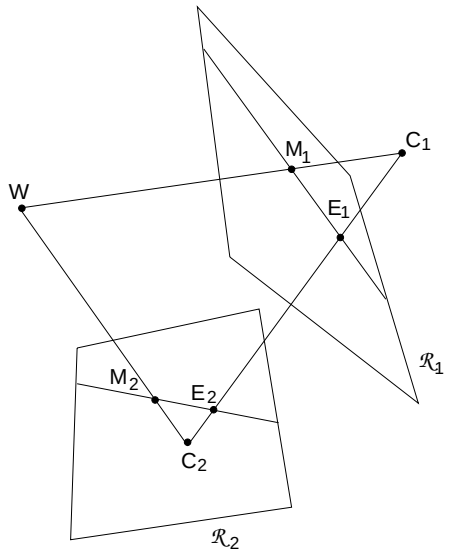

Figure 2: Epipolar geometry. The epipole of the first camera E is the projection of the optical centre C_2 of the second camera (and viceversa).

line is the projection through C_2 of the optical ray of M_1. All the epipolar lines in one image plane pass through a common point (E_1 and E_2 resp.) called the *epipole*, which is the projection of the conjugate optical centre.
When C_1 is in the focal plane of the right camera, the right epipole is at infinity, and the epipolar lines form a bundle of parallel lines in the right image. A very special case is when both epipoles are at infinity, that happens when the line $\mathsf{C}_1\mathsf{C}_2$ (the *baseline*) is contained in both focal planes, i.e., the retinal planes are parallel to the baseline (see Fig. 3). Epipolar lines then form a bundle of parallel lines in both images. Any pair of images can be transformed so that epipolar lines are parallel and horizontal in each image. This procedure is called *rectification*.

3 Rectification of camera matrices

We will assume that the stereo rig is *calibrated*, i.e. the old PPMs $\tilde{\mathbf{P}}_{o1}$ and $\tilde{\mathbf{P}}_{o2}$ are known. This assumption is not strictly necessary [8, 13], but leads to a simpler technique. The idea behind rectification is to define two new perspective matrices $\tilde{\mathbf{P}}_{n1}$ and $\tilde{\mathbf{P}}_{n2}$, which preserve the optical centres and with the baseline contained in the focal planes. This ensures that epipoles are at infinity, hence epipolar lines are parallel. In addition, to have a proper rectification, it is required that epipolar lines are horizontal, and that corresponding points have the same vertical coordinate. We will formalise analytically this requirements in Appendix A, where we will also show that the algorithm given in the present section satisfies that requirements. By now, let the geometric intuition guide us in the development of the rectification algorithm.
The new PPMs will have both the same orientation but different positions. Positions (optical centres) are the same as the old cameras, while orientation changes because we rotate both cameras around the optical centres in such a way that focal planes becomes coplanar and contain the baseline. In order to simplify the algorithm, the rectified PPMs will have also the same intrinsic parameters. The resulting PPMs will differ only in their optical centres. The new camera pair can be thought as a single camera translated along the X axis of its standard reference system. This, intuitively

satisfies the rectification requirements (formal proof in Appendix A).

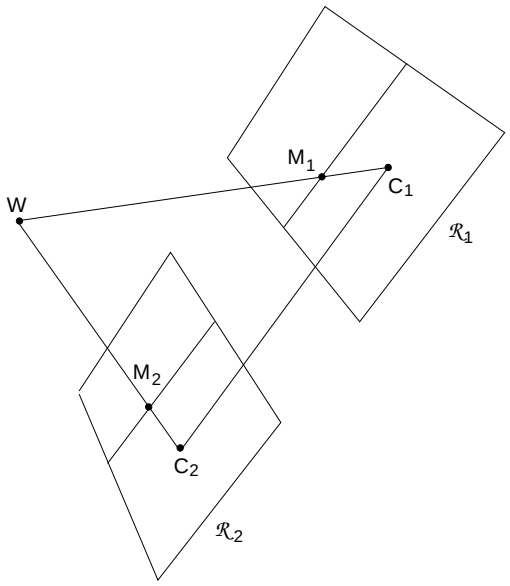

Figure 3: Rectification.Image planes are coplanar and parrallel to the baseline.

Let us think of the new MPPs in terms of their factorisation. From (3) and (9):

$$\tilde{\mathbf{P}}_{n1} = \mathbf{A}[\mathbf{R} \mid -\mathbf{R}\,\mathbf{c}_1] \qquad \tilde{\mathbf{P}}_{n2} = \mathbf{A}[\mathbf{R} \mid -\mathbf{R}\,\mathbf{c}_2] \tag{11}$$

The optical centres $\mathbf{c}_1$ and $\mathbf{c}_2$ are given by the old optical centres, computed with (8). The rotation matrix $\mathbf{R}$ is the same for both PPMs, and is computed as detailed below. The intrinsic parameters matrix $\mathbf{A}$ is also the same for both PPMs, but can be chosen arbitrarily (see MATLAB code). We will specify $\mathbf{R}$ by means of its row vectors

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1^\top \\ \mathbf{r}_2^\top \\ \mathbf{r}_3^\top \end{bmatrix} \tag{12}$$

which are the X, Y andZ axes respectively of the camera standard reference frame, represented in world coordinates.
According to the previous geometric arguments, we take:

1. the new X axis parallel to the baseline: $\mathbf{r}_1 = (\mathbf{c}_1 - \mathbf{c}_2)/||\mathbf{c}_1 - \mathbf{c}_2||$
2. the new Y axis orthogonal to X (mandatory) and to $\mathbf{k}$: $\mathbf{r}_2 = \mathbf{k} \wedge \mathbf{r}_1$
3. the new Z axis orthogonal to XY (mandatory) : $\mathbf{r}_3 = \mathbf{r}_1 \wedge \mathbf{r}_2$

where $\mathbf{k}$ is an arbitrary unit vector, which fixes the position of the new Y axis in the plane orthogonal to X. We take it equal to the Z unit vector of the old left matrix, thereby constraining the new Y axis to be orthogonal to both the new X and the old left Z. The algorithm is given in more details in the MATLAB version, Section 4.

3.1 The rectifying transformation

In order to rectify – let's say – the left image, we need to compute the transformation mapping the image plane of $\tilde{\mathbf{P}}_{o1} = [\mathbf{Q}_{o1}|\tilde{\mathbf{q}}_{o1}]$ onto the image plane of $\tilde{\mathbf{P}}_{n1} = [\mathbf{Q}_{n1}|\tilde{\mathbf{q}}_{n1}]$. We will see that the sought transformation is the collinearity given by the 3×3 matrix $\mathbf{T}_1 = \mathbf{Q}_{n1}\mathbf{Q}_{o1}^{-1}$. The same result will apply to the right image.
For any 3-D point $\mathbf{w}$ we can write

$$\begin{cases} \tilde{\mathbf{m}}_{o1} = \tilde{\mathbf{P}}_{o1}\tilde{\mathbf{w}} \\ \tilde{\mathbf{m}}_{n1} = \tilde{\mathbf{P}}_{n1}\tilde{\mathbf{w}}. \end{cases} \tag{13}$$

According to (10) , the equations of the optical rays are the following (since rectification does not move the optical centre)

$$\begin{cases} \mathbf{w} = \mathbf{c}_1 + \lambda_o \mathbf{Q}_{o1}^{-1}\tilde{\mathbf{m}}_{o1} \\ \mathbf{w} = \mathbf{c}_1 + \lambda_n \mathbf{Q}_{n1}^{-1}\tilde{\mathbf{m}}_{n1}; \end{cases} \tag{14}$$

Hence:

$$\tilde{\mathbf{m}}_{n1} = \lambda \mathbf{Q}_{n1}\mathbf{Q}_{o1}^{-1}\tilde{\mathbf{m}}_{o1}. \tag{15}$$

where λ is an arbitrary scale factor (this is an equality between homogeneous quantities). This is a clearer and more compact result than the one reported in [1], in which $\mathbf{P}_{o1}^{-1}$ is not written as the inverse of a PPM.
The transformation $\mathbf{T}_1$ is then applied to the original left image to produce the rectified image, as in Fig. 6. Note that the pixels (integer-coordinate positions) of the rectified image correspond, in general, to non-integer positions on the original image plane. Therefore, the gray levels of the rectified image are computed by bilinear interpolation.

4 Summary of the algorithm

The process of rectification can be summarised as follows:

- Given a stereo pair of images `I1,I2` and PPMs `Po1,Po2` (obtained by calibration);
- compute `[T1,T2,Pn1,Pn2] = rectify(Po1,Po2)` (see box);
- rectify images by applying `T1` and `T2`.

Reconstruction of 3D position can be performed from the rectified images directly, using `Pn1,Pn2`.

```
function [T1,T2,Pn1,Pn2] = rectify(Po1,Po2)

% RECTIFY: compute  rectification matrices in homogeneous coordinate
%
%         [T1,T2,Pn1,Pn2] = rectify(Po1,Po2)  coomputes the rectified
%         projection matrices "Pn1" and "Pn2", and the transformation
%         of the retinal plane "T1" and  "t2" (in homogeneous coord.)
%         which perform rectification.  The arguments are the two old
%         projection  matrices "Po1" and "Po2".
%         Andrea Fusiello, MVL 1998 (fusiello@dimi.uniud.it)

% factorise old PPMs
[A1,R1,t1] = art(Po1);
[A2,R2,t2] = art(Po1);

% optical centres (unchanged)
c1 = - inv(Po1(:,1:3))*Po1(:,4);
c2 = - inv(Po2(:,1:3))*Po2(:,4);

% new x axis (= direction of the baseline)
v1 = (c1-c2);
% new y axes (orthogonal to new x and old z)
v2 = extp(R1(3,:)',v1);
% new z axes (no choice, orthogonal to baseline and y)
v3 = extp(v1,v2);

% new extrinsic (translation unchanged)
R = [v1'/norm(v1)
     v2'/norm(v2)
     v3'/norm(v3)];

% new intrinsic (arbitrary)
A = (A1 + A2)./2;
A(1,2)=0; % no skew

% new projection matrices
Pn1 = A * [R -R*c1 ];
Pn2 = A * [R -R*c2 ];

% rectifying image transformation
T1 = Pn1(1:3,1:3)* inv(Po1(1:3,1:3));
T2 = Pn2(1:3,1:3)* inv(Po2(1:3,1:3));

------------------------

function [A,R,t] = art(P)
% ART: factorise a PPM as  P=A*[R;t]

Q = inv(P(1:3, 1:3));
[U,B] = qr(Q);

R = inv(U);
t = B*P(1:3,4);
A = inv(B);
A = A ./A(3,3);
```

Given the high diffusion of stereo in research and applications, we have endeavoured to make our algorithm as easily reproducible and usable as possible. To this purpose, we give (see box) the working MATLAB code of the algorithm; the code is simple and compact (20 lines), and the comments enclosed make it understandable without knowledge of MATLAB. Moreover, a "rectification kit" including C and MATLAB code, data sets and user manual can be found on line[2].

5 Experimental Results

We ran tests to verify that the algorithm performed rectification correctly, and also to check that the accuracy of the 3-D reconstruction did not decrease when performed from the rectified images

[2] http://www.dimi.uniud.it/~fusiello.html

directly.

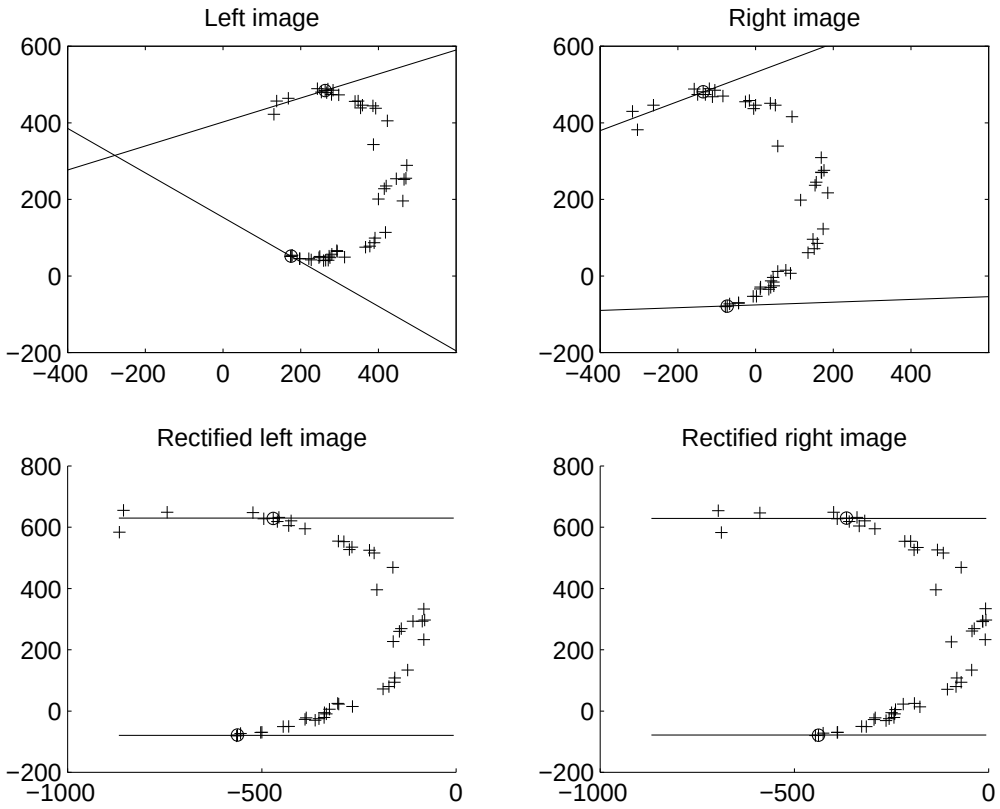

Figure 4: General synthetic stereo pair (top) and rectified pair (bottom). The figure shows the epipolar lines of the points marked with a circle in both images.

5.1 Correctness

The tests used both synthetic and real data. Each set of synthetic data consisted of a cloud of 3-D points and a pair of PPMs; both points and matrices were chosen randomly. For reasons of space, we report only two examples. Figure 5 shows the original and rectified images with a nearly rectified stereo rig: the camera translation was $-[100\ 2\ 3]$ mm and the rotation angles roll=1.5^o, pitch=2^o, yaw=1^o. Figure 4 shows the same with a more general geometry: the camera translation was $-[100\ 20\ 30]$ mm and the rotation angles roll=19^o pitch=32^o and yaw=5^o.
Real-data experiments used calibrated stereo pairs available from the INRIA-Syntim WWW site [14], which include the cameras' PPMs [15]. As for synthetic data, we show the results obtained with a nearly rectified stereo rig (Figure 6) and with a more general stereo geometry (Figure 7). The right image of each pair shows three epipolar lines corresponding to the points marked by a cross in the left image. The pixel coordinates of the rectified images are not constrained to lie in any special part of the image plane, and an arbitrary translation can be applied to both images to bring them in a suitable region of the plane. The output images were cropped to the size of the input images for display purposes only.

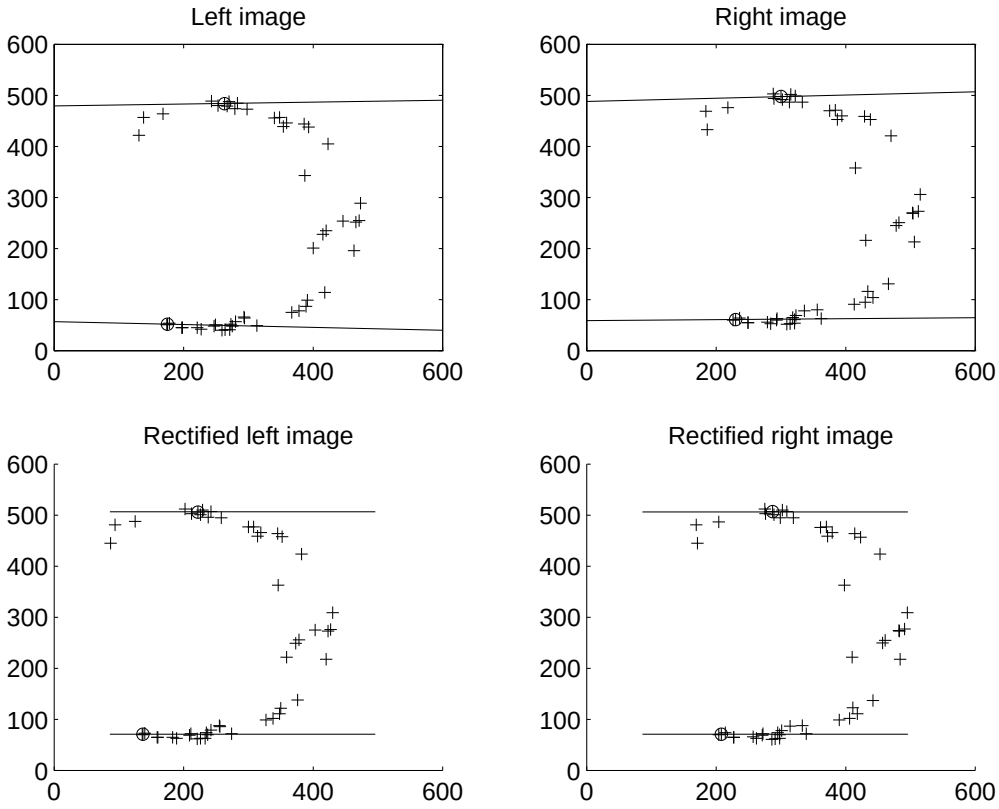

Figure 5: Nearly rectified synthetic stereo pair (top) and rectified pair (bottom). The figure shows the epipolar lines of the points marked with a circle in both images.

Figure 6: Original "Sport" stereo pair (top) and rectified pair (bottom). The right pictures plot the epipolar lines corresponding to the point marked in the left pictures.

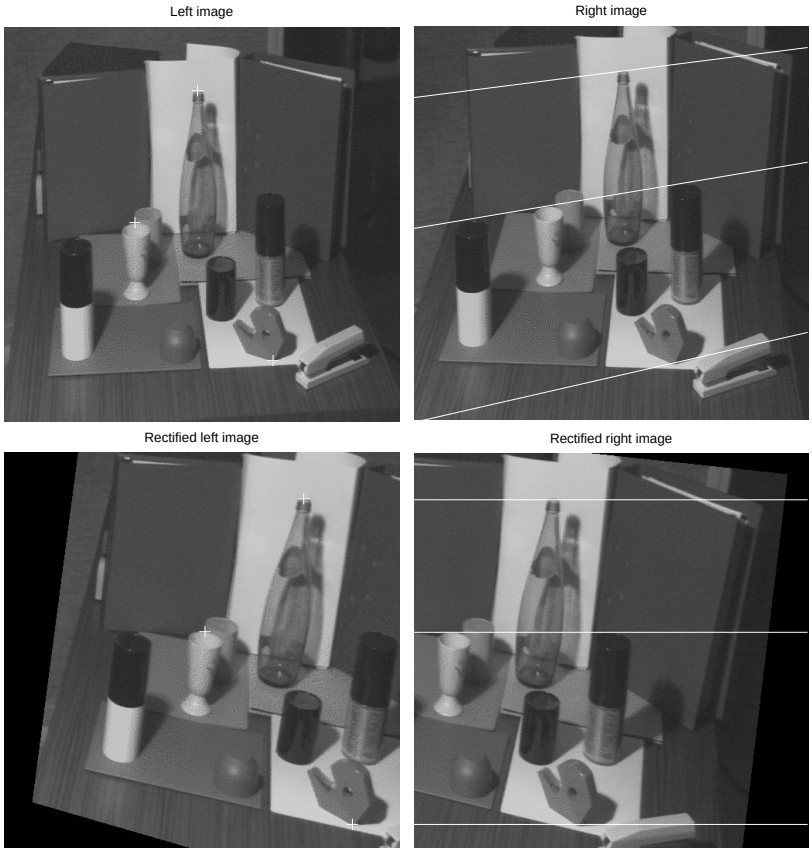

Figure 7: Original "Color" stereo pair (top) and rectified pair (bottom). The right pictures plot the epipolar lines corresponding to the point marked in the left pictures.

5.2 Accuracy

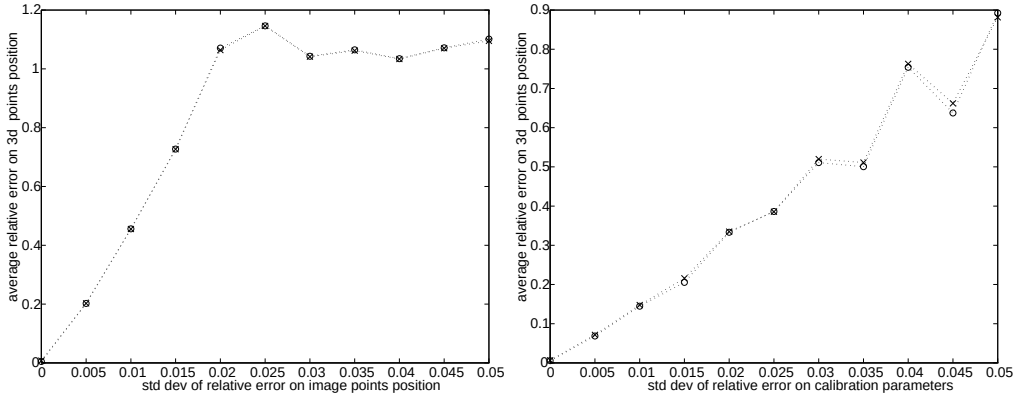

Figure 8: Reconstruction error vs noise levels in the image coordinates (left) and calibration parameters (right) for the stereo rig considered in Figure 2. Crosses refer to reconstruction from rectified images, circles to reconstruction from unrectified images.

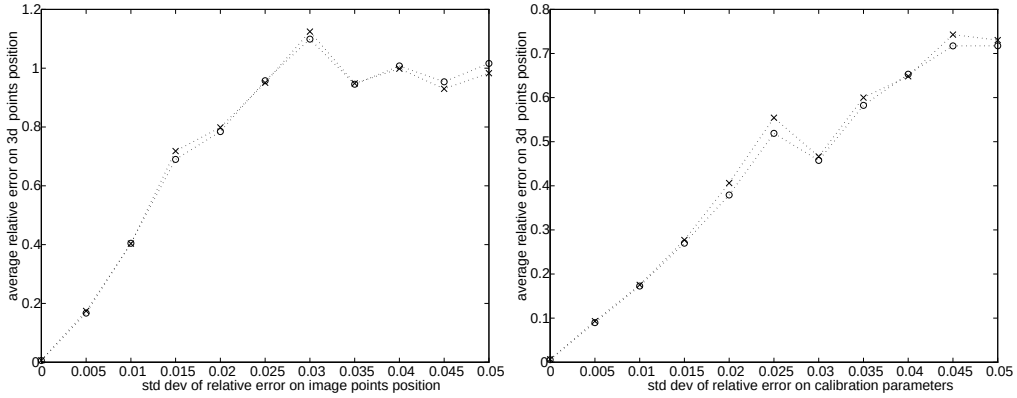

Figure 9: Reconstruction error vs noise levels in the image coordinates (left) and calibration parameters (right) for the stereo rig considered in Figure 3. Crosses refer to reconstruction from rectified images, circles to reconstruction from unrectified images.

In order to evaluate the errors introduced by rectification on reconstruction, we compared the accuracy of 3-D reconstruction computed from original and rectified images. We used synthetic, noisy images of random clouds of 3-D points. Imaging errors were simulated by perturbing the image coordinates, and calibration errors by perturbing the intrinsic and extrinsic parameters, both with additive, Gaussian noise. Reconstruction were performed using a simple linear least-squares algorithm [4]. Figures 8 and 9 show the average (over the set of points) relative error measured on 3-D point position, plotted against noise. Figure 8 shows the results for the stereo rig used in Figure 5, and Figure 9 for the one used in Figure 4. Each point plotted is an average over 100 independent trials. The abscissa is the standard deviation of the relative error on coordinates of image point or calibration parameters.

6 Conclusion

Stereo matching is greatly simplified if the epipolar lines are parallel and horizontal in each image, i.e., if the images are rectified. The correct behaviour of the algorithm has been demonstrated with both synthetic and real images. Interestingly enough, reconstruction can be performed directly from the disparities of the rectified images, using the rectifying PPMs. Our tests show that this process does not introduces appreciable errors compared with reconstructing from the original images. We believe that a general rectification algorithm, together with the material we have made available from the URL given in Section 6, can prove a useful resource for the research and application communities alike.

Acknowledgements

This work benefited from discussions with Bruno Caprile, and was partially supported by grants from the British Council-MURST/CRUI and EPSRC (GR/L18716). The stereo pairs are available from INRIA-Syntim under Copyright.

A Rectification Analysis

In this appendix we will (i) formulate analytically the rectification requirements, and (ii) prove that the our algorithm yield PPMs $\tilde{\mathbf{P}}_{n1}$ and $\tilde{\mathbf{P}}_{n2}$ that satisfies such requirements.
In order for an image pair to be rectified, the epipolar line of the point $\mathbf{m}_1 = (u_1, v_1)^\top$ in the left image should be the horizontal line $v_2 = v_1$ in the right image and viceversa.
In the following, we shall write $\tilde{\mathbf{P}}_{n1}$, $\tilde{\mathbf{P}}_{n2}$ as follows:

$$\tilde{\mathbf{P}}_{n1} = \left[\begin{array}{c|c} \mathbf{s}_1^\top & s_{14} \\ \mathbf{s}_2^\top & s_{24} \\ \mathbf{s}_3^\top & s_{34} \end{array}\right] = [\mathbf{S}|\tilde{\mathbf{s}}] \qquad \tilde{\mathbf{P}}_{n2} = \left[\begin{array}{c|c} \mathbf{d}_1^\top & b_{14} \\ \mathbf{d}_2^\top & b_{24} \\ \mathbf{d}_3^\top & b_{34} \end{array}\right] = [\mathbf{D}|\tilde{\mathbf{d}}]. \tag{16}$$

As we know, the epipolar line of $\tilde{\mathbf{m}}_2$ is the projection of its optical ray onto the left camera, hence its parametric equation writes:

$$\tilde{\mathbf{m}}_1 = \tilde{\mathbf{P}}_{n1}\begin{bmatrix}\mathbf{c}_2 \\ 1\end{bmatrix} + \tilde{\mathbf{P}}_{n1}\begin{bmatrix}\lambda\mathbf{D}^{-1}\tilde{\mathbf{m}}_2 \\ 0\end{bmatrix} = \tilde{\mathbf{e}}_1 + \lambda\mathbf{S}\mathbf{D}^{-1}\tilde{\mathbf{m}}_2 \tag{17}$$

where $\tilde{\mathbf{e}}_1$, the epipole, is the projection of the conjugate optical centre $\mathbf{c}_2$: [3]

$$\tilde{\mathbf{e}}_1 = \tilde{\mathbf{P}}_{n1}\begin{bmatrix}\mathbf{c}_2 \\ 1\end{bmatrix} = \begin{bmatrix}\mathbf{s}_1\mathbf{c}_2 + s_{14} \\ \mathbf{s}_2\mathbf{c}_2 + s_{24} \\ \mathbf{s}_3\mathbf{c}_2 + s_{34}\end{bmatrix}. \tag{18}$$

The parametric equation of the epipolar line of $\tilde{\mathbf{m}}_2$ in image coordinates becomes:

$$\begin{cases} u = [\mathbf{m}_1]_1 = \dfrac{[\tilde{\mathbf{e}}_1]_1 + \lambda[\tilde{\mathbf{n}}]_1}{[\tilde{\mathbf{e}}1]_3 + \lambda[\tilde{\mathbf{n}}]_3} \\ v = [\mathbf{m}_1]_2 = \dfrac{[\tilde{\mathbf{e}}_1]_2 + \lambda[\tilde{\mathbf{n}}]_2}{[\tilde{\mathbf{e}}_1]_3 + \lambda[\tilde{\mathbf{n}}]_3} \end{cases} \tag{19}$$

[3] In this appendix all vector product are scalar product unless otherwise noted

where $\tilde{\mathbf{n}} = \mathbf{S}\mathbf{D}^{-1}\tilde{\mathbf{m}}_2$ and $[.]_i$ is the projection operator extracting the ith component from a vector. Analytically, the direction of each epipolar line can be obtained by taking the derivative of the parametric equations (19) with respect to λ:

$$\begin{cases} \dfrac{du}{d\lambda} = \dfrac{[\tilde{\mathbf{n}}]_1[\tilde{\mathbf{e}}_1]_3 - [\tilde{\mathbf{n}}]_3[\tilde{\mathbf{e}}_1]_1}{([\tilde{\mathbf{e}}_1]_3 + \lambda[\tilde{\mathbf{n}}]_3)^2} \\ \dfrac{dv}{d\lambda} = \dfrac{[\tilde{\mathbf{n}}]_2[\tilde{\mathbf{e}}_1]_3 - [\tilde{\mathbf{n}}]_3[\tilde{\mathbf{e}}_1]_2}{([\tilde{\mathbf{e}}_1]_3 + \lambda[\tilde{\mathbf{n}}]_3)^2} \end{cases}. \tag{20}$$

Note that the denominator is the same in both components, hence it does not affect the direction of the vector. The epipole is rejected to infinity when $[\tilde{\mathbf{e}}_1]_3 = 0$. In this case, the direction of the epipolar lines in the right image doesn't depend on $\mathbf{n}$ any more and all the epipolar lines becomes parallel to vector $[[\tilde{\mathbf{e}}_1]_1\ [\tilde{\mathbf{e}}_1]_2]^\top$. The same holds, *mutatis mutandi*, for the left image .
Hence epipolar lines are are horizontal when:

$$\begin{cases} \mathbf{s}_1\mathbf{c}_2 + s_{14} \neq 0 \\ \mathbf{s}_2\mathbf{c}_2 + s_{24} = 0 \\ \mathbf{s}_3\mathbf{c}_2 + s_{34} = 0 \end{cases} \text{ and } \begin{cases} \mathbf{d}_1\mathbf{c}_1 + d_{14} \neq 0 \\ \mathbf{d}_2\mathbf{c}_1 + d_{24} = 0 \\ \mathbf{d}_3\mathbf{c}_1 + d_{34} = 0 \end{cases} \tag{21}$$

Since the vertical coordinate of the projection of a 3-D point onto the rectifying retinal plane must be the same in both image, we also have:

$$\frac{\mathbf{s}_2\mathbf{w} + s_{24}}{\mathbf{s}_3\mathbf{w} + s_{34}} = \frac{\mathbf{d}_2\mathbf{w} + d_{24}}{\mathbf{d}_3\mathbf{w} + d_{34}}. \tag{22}$$

Therefore, a pair of PPMs are rectified as soon as (21) and (22) hold. Let us check that this is true for $\tilde{\mathbf{P}}_{n1}$ and $\tilde{\mathbf{P}}_{n2}$.
From (11) we obtain

$$\begin{array}{lll} s_{14} = -\mathbf{s}_1\mathbf{c}_1 & d_{14} = -\mathbf{d}_1\mathbf{c}_2 & \mathbf{s}_1 = \mathbf{d}_1 \\ s_{24} = -\mathbf{s}_2\mathbf{c}_1 & d_{24} = -\mathbf{d}_2\mathbf{c}_2 & \mathbf{s}_2 = \mathbf{d}_2 \\ s_{34} = -\mathbf{s}_3\mathbf{c}_1 & d_{34} = -\mathbf{d}_3\mathbf{c}_2 & \mathbf{s}_3 = \mathbf{d}_3 \end{array} \tag{23}$$

From the factorisation (3), assuming $\gamma = 0$, we obtain

$$\begin{bmatrix} \mathbf{s}_1^\top \\ \mathbf{s}_2^\top \\ \mathbf{s}_3^\top \end{bmatrix} = \mathbf{A}\mathbf{R} = \begin{bmatrix} \alpha_u\mathbf{r}_1^\top + u_0\mathbf{r}_3^\top \\ \alpha_v\mathbf{r}_2^\top + v_0\mathbf{r}_3^\top \\ \mathbf{r}_3^\top \end{bmatrix} \tag{24}$$

From the construction of $\mathbf{R}$, we have that $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{r}_3$ are mutually orthogonal and $\mathbf{r}_1 = \beta(\mathbf{c}_1 - \mathbf{c}_2)$ with $\beta = 1/||\mathbf{c}_1 - \mathbf{c}_2||$.
From all this facts, the following four identity follow:

$$\mathbf{s}_1(\mathbf{c}_1 - \mathbf{c}_2) = \beta\mathbf{s}_1\mathbf{r}_1 = \beta(\alpha_u\mathbf{r}_1 + u_0\mathbf{r}_3)\mathbf{r}_1 = \beta(\alpha_u\mathbf{r}_1\mathbf{r}_1 + u_0\mathbf{r}_3\mathbf{r}_1) = \beta\alpha_u \neq 0 \tag{25}$$

$$\mathbf{s}_2(\mathbf{c}_1 - \mathbf{c}_2) = \beta\mathbf{s}_2\mathbf{r}_1 = \beta(\alpha_v\mathbf{r}_2 + v_0\mathbf{r}_3)\mathbf{r}_1 = \beta(\alpha_v\mathbf{r}_2\mathbf{r}_1 + v_0\mathbf{r}_3\mathbf{r}_1) = 0 \tag{26}$$

$$\mathbf{s}_3(\mathbf{c}_1 - \mathbf{c}_2) = \beta\mathbf{s}_3\mathbf{r}_1 = \beta\mathbf{r}_3\mathbf{r}_1 = 0 \tag{27}$$

$$\mathbf{s}_2 \wedge \mathbf{s}_3 = \mathbf{s}_2 \wedge \mathbf{r}_3 = \lambda(\mathbf{r}_2 \wedge \mathbf{r}_3) = \lambda\mathbf{r}_1 \tag{28}$$

In (28) λ is a suitable scalar, which comes from the fact that $\mathbf{s}_2$ is a linear combination of $\mathbf{r}_2$ and $\mathbf{r}_3$.
Equation (21) follows easily from (25) (26)(27)
Equation (22) is equivalent to

$$(\mathbf{s}_2\mathbf{w} + s_{24})(\mathbf{d}_3\mathbf{w} + d_{34}) = (\mathbf{s}_3\mathbf{w} + s_{34})(\mathbf{d}_2\mathbf{w} + d_{24}). \tag{29}$$

Expanding:

$$\begin{aligned} -\mathbf{s}_2(\mathbf{c}_1 - \mathbf{c}_2)\mathbf{s}_3\mathbf{w} + (\mathbf{s}_2\mathbf{c}_1)(\mathbf{s}_3\mathbf{c}_2) - (\mathbf{s}_2\mathbf{c}_2)(\mathbf{s}_3\mathbf{c}_1) = (\mathbf{s}_2\mathbf{c}_1)(\mathbf{s}_3\mathbf{c}_2) - (\mathbf{s}_2\mathbf{c}_2)(\mathbf{s}_3\mathbf{c}_1) = \\ (\mathbf{s}_2 \wedge \mathbf{s}_3)(\mathbf{c}_1 \wedge \mathbf{c}_2) = \lambda\mathbf{r}_1(\mathbf{c}_1 \wedge \mathbf{c}_2) = \lambda\beta(\mathbf{c}_1 - \mathbf{c}_2)(\mathbf{c}_1 \wedge \mathbf{c}_2) = 0 \end{aligned} \tag{30}$$

using (26)(28) and properties of the external product.

References

[1] N. Ayache. *Artificial Vision for Mobile Robots: Stereo Vision and Multisensory Perception.* The MIT Press, 1991.

[2] B. Caprile and V. Torre. Using vanishing points for camera calibration. *International Journal of Computer Vision*, 4:127–140, 1990.

[3] I. J. Cox, S. Hingorani, B. M. Maggs, and S. B. Rao. A maximum likelihood stereo algorithm. *Computer Vision and Image Understanding*, 63(3):542–567, May 1996.

[4] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint.* The MIT Press, Cambridge, 1993.

[5] O. Faugeras and G. Toscani. Camera calibration for 3D computer vision. In *Proceedings of the International Workshop on Machine Vision and Machine Intelligence*, Tokyo, Japan, February 1987.

[6] A. Fusiello, V. Roberto, and E. Trucco. Efficient stereo with multiple windowing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 858–863, Puerto Rico, June 1997. IEEE Computer Society Press.

[7] D. Geiger, B. Ladendorf, and A. Yuille. Occlusions and binocular stereo. *International Journal of Computer Vision*, 14(3):211–226, April 1995.

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

[9] S. S. Intille and A. F. Bobick. Disparity-space images and large occlusion stereo. In Jan-Olof Eklundh, editor, *European Conference on Computer Vision*, pages 179–186, Stockholm, Sweden, May 1994. Springer-Verlag.

[10] T. Kanade and M. Okutomi. A stereo matching algorithm with an adaptive window: Theory and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9):920–932, September 1994.

[11] D. V. Papadimitriou and T. J. Dennis. Epipolar line estimation and rectification for stereo images pairs. *IEEE Transactions on Image Processing*, 3(4):672–676, April 1996.

[12] L Robert. Camera calibration without feature extraction. *Computer Vision, Graphics, and Image Processing*, 63(2):314–325, March 1995. also INRIA Technical Report 2204.

[13] L. Robert, C. Zeller, O. Faugeras, and M. Hébert. Applications of non-metric vision to some visually-guided robotics tasks. In Y. Aloimonos, editor, *Visual Navigation: From Biological Systems to Unmanned Ground Vehicles*, chapter 5, pages 89–134. Lawrence Erlbaum Associates, 1997.

[14] Stereograms of the syntim project, Copyright Institut National de Recherche en Informatique et Automatique, 1994, 1995, 1996. http://www-syntim.inria.fr/syntim/analyse/paires-eng.html.

[15] Jean-Philippe Tarel and André Gagalowicz. Calibration de caméra à base d'ellipses. *Traitement du Signal*, 12(2):177–187, 1995.

[16] C. Tomasi and R. Manduchi. Stereo without search. In B. Buxton and R. Cipolla, editors, *European Conference on Computer Vision*, pages 452–465, Cambridge (UK), April 1996.

[17] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, August 1987.