Rectifying homographies for stereo vision: analytical solution for minimal distortion

Pasquale Lafiosca[a]   Marta Ceccaroni[b]

[a]Integrated Vehicle Health Management Centre, Cranfield University, United Kingdom
[b]School of Aerospace, Cranfield University, United Kingdom

Abstract

Stereo rectification is the determination of two image transformations (or *homographies*) that map *corresponding points* on the two images, projections of the same point in the 3D space, onto the same horizontal line in the transformed images. Rectification is used to simplify the subsequent *stereo correspondence* problem and speeding up the matching process. Rectifying transformations, in general, introduce perspective distortion on the obtained images, which shall be minimised to improve the accuracy of the following algorithm dealing with the stereo correspondence problem. The search for the optimal transformations is usually carried out relying on numerical optimisation. This work proposes a closed-form solution for the rectifying homographies that minimise perspective distortion. The experimental comparison confirms its capability to solve the convergence issues of the previous formulation. Its Python implementation is provided.

Keywords: Stereo vision, Stereo image processing, Image rectification, Epipolar geometry

1 Introduction

Stereo vision gained a prominent role among Computer Vision technologies as it allows machines to perceive depth. Thanks to pinhole camera model, calibration and epipolar geometry, a considerable simplification can be performed before attempting to solve the stereo correspondence. Among these simplifications, rectification is almost always conducted to obtain horizontal and aligned epipolar lines, so that the following stereo matching algorithm can work along the x-axis only, with significant performance increase.

However, rectification transformations generally introduce distortion, which, in turn, impairs the performance of the following stereo matching algorithm. For this reason, minimal-distortion algorithms are proposed in the literature. In this paper the closed-form solution of the rectifying homographies that minimise perspective distortion is derived by means of the metric introduced by Loop and Zhang [1]. This is applied to a calibrated stereo rig, where all the intrinsic and extrinsic parameters are known.

The hereby presented formulation of the minimising solution is enabled by a new geometrical interpretation of the problem, which simplifies the distortion

metric expression. The solution found is general and, therefore, valid for every relative position of a given stereo rig, even in extreme configurations and when the previous algorithm [1] fails to provide an initial guess for minimisation (see Sec. 5). Furthermore, the proposed method overcomes the need for optimisation libraries and is thus computationally more efficient by avoiding numerical minimisation loops.

1.1 Background

Emulating the human vision system, stereo vision derives 3D information by comparison of 2D digital images taken from two distinct camera positions, usually known from calibration [2]. Given a stereo image pair, the third coordinate of each object in the scene is extracted by first solving the *stereo correspondence* problem [3, 4], namely the problem of locating the projections of that object in the two images, and then applying *triangulation* [5] to recover the position of the object in the 3D world.

Taking advantage of the geometric relations between the 3D points and their projections onto the 2D images, the so-called *epipolar constraint*, the stereo correspondence problem can be reduced to one dimension. This means that the search of each pair of corresponding points is thus carried on along one, usually oblique, *epipolar* line in each picture. To further simplify the problem, *image rectifications* can be applied, transforming the stereo correspondence problem into a search along a horizontal line, with a consequent, significant improvement in efficiency.

Image rectifications are a family of 2D transformations that can be applied to a couple of non-coplanar stereo images to re-project them onto a *rectifying plane* such that, on the transformed images, corresponding points will lie on the same horizontal line. The drawback of rectification is that it introduces a certain amount of distortion in the resulting images, that decreases the accuracy of the subsequent stereo matching algorithms. Hence, the importance of identifying, among the rectifying transformations in the family, the ones that minimise distortion. Hereafter, in continuity with the previous algorithm [1], we will refer to *perspective distortion* simply as distortion.

This paper make the rectifications minimising distortion explicit, finding their analytic formulation via a geometrical derivation. The basic notation and metric are given in Sec. 2. Our novelties, the geometrical interpretation and derivation of the closed form solution, are reported in Sec. 3. The main steps of the algorithm are listed in Sec. 4, followed by a discussion and examples in Sec. 5. Conclusions follow in Sec. 6. The Python 3 implementation of the algorithm is also provided.

1.2 Previous work

Several algorithms for stereo image rectification are available in the literature. Among the mostly accepted and applied, the one by Fusiello et al. [6] essentially fixes the rectifying plane using the sole orientation of one of the two cameras. Despite being a simple and compact, this algorithm is based on an arbitrary non-optimal choice of one rectifying plane (see Sec. 5.1). This choice leads, for highly skew configurations of the cameras, to the introduction of a massive amount of

distortion in the transformed images, with the consequence of worsening the performances of the algorithm dealing with the stereo correspondence.

The importance of minimising the distortion introduced by the rectifying transformations is highlighted by Loop and Zhang [1]. In their work, the authors decompose rectifying homographies as a combination of a *perspective*, a *similarity* and a *shearing* sub-transformations, and minimise the global distortion. A metric for measuring perspective distortion is thus presented. Finally, using an iterative, numerical procedure, the pair of homographies introducing the minimum perspective distortion is found. However, the initial guess thereby suggested cannot always be found [7].

With the aim to minimise distortions at pixel level, a different distortion metric was proposed by Gluckman and Nayar [8], making use of the Jacobian of the transformation to track changes in local image areas. This method proposes an approximate solution or requires to follow a complex iterative procedure to obtain the optimal one.

A first-attempt of a closed-form solution was proposed by Sun [9]. Here the rectification transformations are readily found estimating the fundamental matrix. However, it leads to a particular solution without the criteria of reducing perspective distortion.

Lately the attention of researchers has been focused on strategies for rectifying *uncalibrated* stereo rigs [10, 11, 12, 13, 14]. These methods do not require a calibration process and try to estimate camera parameters using the scene itself. For example, a handy solution for uncalibrated dual lens cameras [12] relies on key-points that have to be matched in the images. This approach is based on the very limiting assumption of a *small-drift* between the cameras poses. Obviously this does not apply to general camera poses, like verging cameras, and calibration is still important to reach high accuracy and efficiency levels. Different approaches focus on rectifying fisheye images [15, 16, 17]. Uncalibrated and fisheye rectification algorithms are outside the scope of this article. To date, the vast majority of stereo vision systems employs calibration, that allows for a *metric* reconstruction with the subsequent triangulation.

2   Setting the Problem

This Section summarises the work of Loop and Zhang [1] to provide the basis for the following closed-form solution.

2.1   Pinhole Cameras and Epipolar Geometry

Let us consider a calibrated stereo rig composed of two *pinhole* cameras, where distortion caused by lens imperfections has been already corrected. As most definitions hereafter will be analogous for both cameras, we will define them only for one camera (definitions for the second camera will be obtained replacing the subscript with 2). Let $\mathbf{A}_1 \in \mathbb{R}^{3\times3}$ be the intrinsic matrix of the left camera, with $\mathbf{R}_1 \in \mathbb{R}^{3\times3}$ and $\mathbf{t}_1 \in \mathbb{R}^3$ its rotation matrix and translation vector, respectively, describing the position of the first *Camera Coordinate System* (CCS) with respect to *World Coordinate System* (WCS), as represented in Fig. 1, with a slight abuse of notation for axes names to ease visualisation.
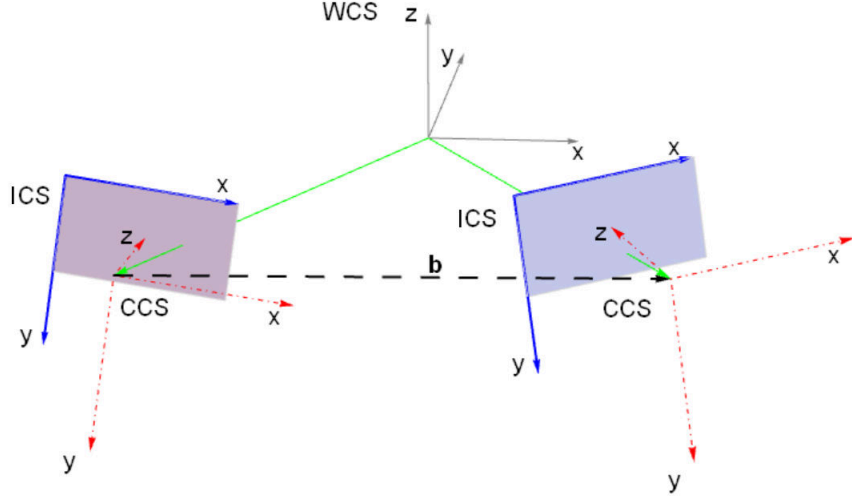
Figure 1: Representation of coordinate systems. CCSs are in red, dot-dashed. ICSs are in blue. WCS is in grey and the baseline is dashed black.

Call $\mathbf{o}_1 = -\mathbf{R}_1^{-1} \cdot \mathbf{t}_1$ the position of the optical center in the WCS. Hereafter, unless otherwise stated, elements are expressed in WCS. The *baseline* is the vector $\mathbf{b} = \mathbf{o}_2 - \mathbf{o}_1$ going from the first to the second camera center. Additionally, an *Image Coordinate System* (ICS) is defined on the image plane of each camera, where the left image I_1 forms (located at $z = 1$ in CCS) with the x and y axes parallel to the respective CCS ones and origin in the upper left corner of the image, corresponding to $\left[-\frac{[\mathbf{A}_1]_{13}}{[\mathbf{A}_1]_{11}} \quad -\frac{[\mathbf{A}_1]_{23}}{[\mathbf{A}_1]_{22}} \quad 1 \right]$ in CCS. The notation $[\mathbf{A}_1]_{ij}$ indicates the (i, j) element of $\mathbf{A}_1$. The ICS system is shown in Fig. 1 as well.

Given two corresponding image points $\mathbf{p}_1 \in I_1$ and $\mathbf{p}_2 \in I_2$, each one in its ICS and expressed as homogeneous coordinates (i.e. $\mathbf{p}_1 \in \mathbb{R}^3$, with unitary 3rd coordinate), the epipolar geometry constraint is defined as:

$$\mathbf{p}_2^T \cdot \mathbf{F} \cdot \mathbf{p}_1 = 0 \tag{1}$$

where $\mathbf{F}$ is the fundamental matrix, a 3×3 matrix of rank 2 that can only be determined up to a scale factor [3, 18], assumed as known. $\mathbf{F}$ maps a point on I_1 to a *line* on I_2, and vice versa (using its transpose). Given that we work in a *projective space*, all the points are defined up to an arbitrary scaling factor.

On each image all the epipolar lines will intersect in a single point called *epipole*. Let $\mathbf{e}_1 \in I_1$ and $\mathbf{e}_2 \in I_2$ be the two epipoles in homogeneous ICS, they can be defined as the left and right *kernels* of $\mathbf{F}$, so that:

$$\mathbf{F} \cdot \mathbf{e}_1 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T = \mathbf{F}^T \cdot \mathbf{e}_2 \tag{2}$$

Geometrically, $\mathbf{e}_1$ is the projection of $\mathbf{o}_2$ on the image I_1. Similarly for $\mathbf{e}_2$. It is worth noticing that each pair of corresponding epipolar lines lies on a same plane together with the baseline.

2.2 Rectification

Rectification is the problem of determining two homographies that map corresponding epipolar lines onto parallel horizontal lines sharing the same y-coordinate (i.e. sending the epipoles to ∞). Thus, rectified images must have the new fundamental matrix in the form [3]:

$$\overline{\mathbf{F}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \tag{3}$$

Defining the rectified points $\overline{\mathbf{p}}_1 = \mathbf{H}_1 \cdot \mathbf{p}_1$ and $\overline{\mathbf{p}}_2 = \mathbf{H}_2 \cdot \mathbf{p}_2$, then, from Eq. (1), we get:

$$\begin{aligned} \mathbf{p}_2^T \cdot \mathbf{F} \cdot \mathbf{p}_1 &= \overline{\mathbf{p}}_2^T \cdot \mathbf{H}_2^{-T} \cdot \mathbf{F} \cdot \mathbf{H}_1^{-1} \cdot \overline{\mathbf{p}}_1 \\ &= \overline{\mathbf{p}}_2^T \cdot \overline{\mathbf{F}} \cdot \overline{\mathbf{p}}_1 \\ &= 0 \end{aligned} \tag{4}$$

where $\mathbf{H}_1$ and $\mathbf{H}_2$ are the left and right sought-after homographies, respectively.

2.3 Perspective Distortion

Following the original procedure [1], given a generic homography $\mathbf{H}_1$, we scale it by its last element and decompose as:

$$\mathbf{H}_1 = \begin{bmatrix} u_{1a} & u_{1b} & u_{1c} \\ v_{1a} & v_{1b} & v_{1c} \\ w_{1a} & w_{1b} & 1 \end{bmatrix} = \mathbf{H}_{1a} \cdot \mathbf{H}_{1p} \tag{5}$$

where $\mathbf{H}_{1a}$ is an affine transformation and $\mathbf{H}_{1p}$ is a purely perspective transformation in the form:

$$\mathbf{H}_{1p} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ w_{1a} & w_{1b} & 1 \end{bmatrix} \tag{6}$$

$\mathbf{H}_{1p}$ is the sole responsible for introducing perspective distortion. The affine transformation will then be:

$$\mathbf{H}_{1a} = \begin{bmatrix} u_{1a} - u_{1c}w_{1a} & u_{1b} - u_{1c}w_{1b} & u_{1c} \\ v_{1a} - v_{1c}w_{1a} & v_{1b} - v_{1c}w_{1b} & v_{1c} \\ 0 & 0 & 1 \end{bmatrix} \tag{7}$$

$\mathbf{H}_{1a}$ can be decomposed further in a shearing transformation, followed by a similarity transformation [1]. The same decomposition is derived for $\mathbf{H}_2$.

Let $\mathbf{p}_1 = [x_1 \ y_1 \ 1]^T$ be a generic point on I_1, in homogeneous ICS, then the

perspective transformed point is:

$$
\begin{aligned}
\mathbf{p}'_1 &= \mathbf{H}_{1p} \cdot \mathbf{p}_1 \\
&= \begin{bmatrix} x_1 \\ y_1 \\ w_{1a}x_1 + w_{1b}y_1 + 1 \end{bmatrix} \\
&= \begin{bmatrix} x_1 \\ y_1 \\ \mathbf{w}_1^T \cdot \mathbf{p}_1 \end{bmatrix} \\
&\propto \begin{bmatrix} \frac{x_1}{\mathbf{w}_1^T \cdot \mathbf{p}_1} \\ \frac{y_1}{\mathbf{w}_1^T \cdot \mathbf{p}_1} \\ 1 \end{bmatrix}
\end{aligned}
\tag{8}
$$

where $\mathbf{w}_1^T = \begin{bmatrix} w_{1a} & w_{1b} & 1 \end{bmatrix}^T$.

The codomain of I_1 through $\mathbf{H}_{1p}$ must be intended as including the point ∞, as per the *hyperplane model* defined to handle perspective transformations [19].

Notice that if $w_{1a} = w_{1b} = 0$ there is no perspective component in the rectifying transformations, then $\mathbf{H}_1$ is a purely affine transformation. If this is true also for $\mathbf{H}_2$, then the stereo rig is perfectly frontoparallel and the image pair is already rectified. However this, in general, does not happen.

2.4   Distortion Metric

We refer to the distortion metric introduced by Loop and Zhang [1]. The aim is to select $\mathbf{H}_1$ and $\mathbf{H}_2$ "as affine as possible", meaning that the elements w_{1a}, w_{1b}, w_{2a} and w_{2b} should be chosen so to introduce less distortion.

Taking the average point $\mathbf{p}_{1c} = \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{p}_{1i}$ as reference, with n_1 total number of pixels of I_1, the amount of distortion on I_1 is defined as:

$$
\sum_{i=1}^{n_1} \frac{\mathbf{w}_1^T \cdot (\mathbf{p}_{1i} - \mathbf{p}_{1c})}{\mathbf{w}_1^T \cdot \mathbf{p}_{1c}}
\tag{9}
$$

The goal is to find the global minimum of the sum of the distortion of both images:

$$
\sum_{i=1}^{n_1} \frac{\mathbf{w}_1^T \cdot (\mathbf{p}_{1i} - \mathbf{p}_{1c})}{\mathbf{w}_1^T \cdot \mathbf{p}_{1c}} + \sum_{i=1}^{n_2} \frac{\mathbf{w}_2^T \cdot (\mathbf{p}_{2i} - \mathbf{p}_{2c})}{\mathbf{w}_2^T \cdot \mathbf{p}_{2c}}
\tag{10}
$$

Then (10) can be rewritten in matrix form as:

$$
\frac{\mathbf{w}_1^T \cdot \mathbf{P}_1 \cdot \mathbf{P}_1^T \cdot \mathbf{w}_1}{\mathbf{w}_1^T \cdot \mathbf{P}_{c1} \cdot \mathbf{P}_{c1}^T \cdot \mathbf{w}_1} + \frac{\mathbf{w}_2^T \cdot \mathbf{P}_2 \cdot \mathbf{P}_2^T \cdot \mathbf{w}_2}{\mathbf{w}_2^T \cdot \mathbf{P}_{c2} \cdot \mathbf{P}_{c2}^T \cdot \mathbf{w}_2}
\tag{11}
$$

where:

$$
\mathbf{P}_1 \cdot \mathbf{P}_1^T = \frac{wh}{12} \begin{bmatrix} w^2 - 1 & 0 & 0 \\ 0 & h^2 - 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\tag{12}
$$

and:

$$
\mathbf{P}_{c1} \cdot \mathbf{P}_{c1}^T = \begin{bmatrix} \frac{(w-1)^2}{4} & \frac{(w-1)(h-1)}{4} & \frac{(w-1)}{2} \\ \frac{(w-1)(h-1)}{4} & \frac{(h-1)^2}{4} & \frac{(h-1)}{2} \\ \frac{(w-1)}{2} & \frac{(h-1)}{2} & 1 \end{bmatrix}
\tag{13}
$$

where w and h are pixel width and height of I_1. Similarly applies to I_2.

3 Geometric Interpretation and Analytical Derivation of the Minimum

A geometric interpretation of the family of rectifying homographies is introduced hereafter, which paves the way to find the analytical formulation for the global minimum of Eq. 10 and build the corresponding rectifying homographies that minimise perspective distortion.

3.1 Geometric Interpretation

It is known that finding rectifying transformations can alternatively be seen as determining a new *common orientation* for a pair of novel *virtual* cameras projecting the images on the same principal plane, such that epipolar lines become horizontal (Fig. 2).

In this subsection we will demonstrate that this can only be achieved if the $\hat{x}$-axis of the new *common orientation* is chosen parallel to the baseline, while the $\hat{z}$-axis can be arbitrarily chosen, thus generating the full rectifying family.
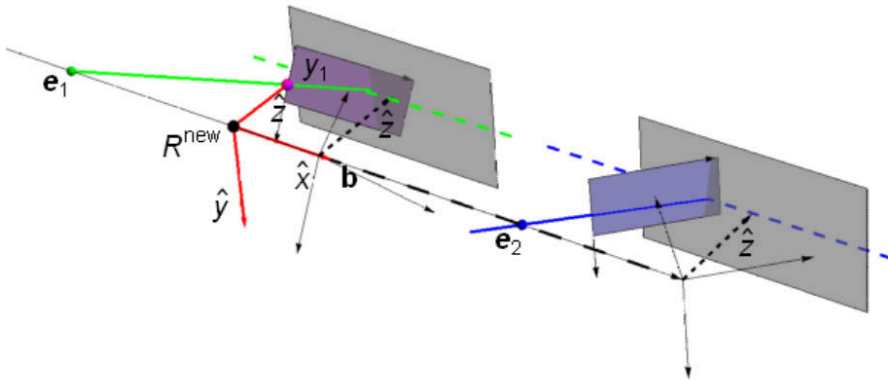

Figure 2: Common orientation of the virtual camera pair (red), projecting on a common plane (gray). $\hat{\mathbf{x}}$ is parallel to the baseline $\mathbf{b}$ (black, dashed). The corresponding epipolar lines (blue, green) and $\hat{\mathbf{z}}$ are identified by y_1 (magenta). Rectified epipolar lines are green and blue, dashed.

The new common orientation will thus be defined as:

$$\mathbf{R}^{new} = \begin{bmatrix} \hat{\mathbf{x}}^T \\ \hat{\mathbf{y}}^T \\ \hat{\mathbf{z}}^T \end{bmatrix} \tag{14}$$

where $\hat{\mathbf{x}} = \frac{\mathbf{b}}{\|\mathbf{b}\|}$ (the versor of the baseline) while $\hat{\mathbf{z}}$ and $\hat{\mathbf{y}}$ can be chosen accordingly to form a Cartesian coordinate system.

Each homography will have to cancel the corresponding intrinsic camera matrix and camera rotation, re-orient the camera with the new chosen orientation and apply another *affine* transformation, namely:

$$\begin{aligned} \mathbf{H}_1 &= \mathbf{K}_1 \cdot \mathbf{R}^{new} \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \\ \mathbf{H}_2 &= \mathbf{K}_2 \cdot \mathbf{R}^{new} \cdot (\mathbf{A}_2 \cdot \mathbf{R}_2)^{-1} \end{aligned} \tag{15}$$

where $\mathbf{K}_1, \mathbf{K}_2 \in \mathbb{R}^{3\times3}$ are arbitrary *affine* transformations.

The geometric interpretation above is based on the following Theorem.

Theorem 1. *The epipolar lines of two cameras are corresponding horizontal lines if and only if the two cameras share the same orientation, with the $\hat{x}$-axis parallel to their baseline.*

To prove this theorem we first introduce two Lemmas.

Theorem 1. *The fundamental matrix $\mathbf{F}$, with an abuse of notation, can be written as:*

$$\mathbf{F} \propto \mathbf{G} \times \mathcal{H} = \begin{bmatrix} [\mathbf{G}]_2 [\mathcal{H}]_{3*} - [\mathbf{G}]_3 [\mathcal{H}]_{2*} \\ [\mathbf{G}]_3 [\mathcal{H}]_{1*} - [\mathbf{G}]_1 [\mathcal{H}]_{3*} \\ [\mathbf{G}]_1 [\mathcal{H}]_{2*} - [\mathbf{G}]_2 [\mathcal{H}]_{1*} \end{bmatrix} \tag{16}$$

where $\mathbf{G} = \mathbf{A}_2 \cdot \mathbf{R}_2 \cdot \mathbf{b} \in \mathbb{R}^3$ and $\mathcal{H} = \mathbf{A}_2 \cdot \mathbf{R}_2 \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \in \mathbb{R}^{3\times3}$. The ith row of a matrix is denoted as $[\ ]_{i*}$.

Proof (Lemma 1): let be $\mathbf{X} \in \mathbb{R}^3$ a point in WCS and $\mathbf{x}_1 \in I_1$ and $\mathbf{x}_2 \in I_2$ its projections in ICS, then $\mathbf{X}$ is a non-trivial solution of the linear system [20]:

$$\begin{cases} \mathbf{A}_1 \cdot (\mathbf{R}_1 \cdot \mathbf{X} + \mathbf{t}_1) = \lambda \mathbf{x_1} \\ \mathbf{A}_2 \cdot (\mathbf{R}_2 \cdot \mathbf{X} + \mathbf{t}_2) = \mu \mathbf{x_2} \end{cases} \tag{17}$$

where $\lambda, \mu \in \mathbb{R}$. Setting $\widetilde{\mathbf{X}} = \mathbf{A}_1 \cdot (\mathbf{R}_1 \cdot \mathbf{X} + \mathbf{t}_1)$, yields:

$$\begin{cases} \mathbb{I}_3 \cdot \widetilde{\mathbf{X}} & = \lambda \mathbf{x_1} \\ \mathcal{H} \cdot \widetilde{\mathbf{X}} - \mathbf{G} & = \mu \mathbf{x_2} \end{cases} \tag{18}$$

where $\mathbb{I}_3 \in \mathbb{R}^{3\times3}$ is the identity matrix. Therefore:

$$Det\left(\begin{bmatrix} \mathbb{I}_3 & \mathbf{0} & \mathbf{x}_1 & \mathbf{0} \\ \mathcal{H} & \mathbf{G} & \mathbf{0} & \mathbf{x}_2 \end{bmatrix} \right) = 0 \tag{19}$$

which is linear in both $\mathbf{x}_1$ and $\mathbf{x}_2$ and can thus be rewritten as:

$$\mathbf{x}_2^T \cdot \mathbf{F} \cdot \mathbf{x}_1 = 0 \tag{20}$$

with:

$$[\mathbf{F}]_{ij} = (-1)^{i+j} Det(\hat{\mathcal{Q}}^{i,j}), \qquad \forall i, j = 1, 2, 3. \tag{21}$$

where $\hat{\mathcal{Q}}^{i,j} \in \mathbb{R}^{4\times4}$ is equal to the matrix $\mathcal{Q} = \begin{bmatrix} \mathbb{I}_3 & \mathbf{0} \\ \mathcal{H} & \mathbf{G} \end{bmatrix} \in \mathbb{R}^{6\times6}$ with the ith and $(3+j)^{th}$ rows dropped. Furthermore:

$$[\mathbf{F}]_{ij} = (-1)^{i+1} Det(\bar{\mathcal{S}}^{i,j}) \tag{22}$$

with the first column of $\bar{\mathcal{S}}^{i,j} \in \mathbb{R}^{2 \times 2}$ equal to the jth column of the matrix $\mathcal{H}$ with the ith row dropped, and the second column equal to $\mathbf{G}$ with the ith element dropped. Then Eq. (22) is equivalent to Eq. (16) $\square$.

Remark: it can be shown that the matrix $\mathcal{H}$ is the matrix representing the axes of the first CCS as seen from the second CCS and that $\mathbf{e}_2 = \frac{\mathbf{G}}{\|\mathbf{G}\|}$. Therefore the fundamental matrix can be expressed as $\mathbf{F} \propto \mathbf{e}_2 \times \mathcal{H}$, similarly to Eq. (16).

Theorem 2. *The epipolar lines of the second camera are horizontal if and only if the x-axis of the second camera is parallel to the baseline.*

Proof (Lemma 2): by the definition of the fundamental matrix (see Sec 2.1), the epipolar lines of the second camera are horizontal *if and only if* the first row of the fundamental matrix $\mathbf{F}$ is null. Using Eq. *Lemma 1* the first row of $\mathbf{F}$ can be expressed as:

$$[\mathbf{G}]_2 [\mathcal{H}]_{3*} - [\mathbf{G}]_3 [\mathcal{H}]_{2*} \tag{23}$$

From the remark above, the rows of $\mathcal{H}$ form a base of linearly independent vectors, therefore the only linear combination that can set the formula (23) to zero is the trivial one, with $[\mathbf{G}]_2 = [\mathbf{G}]_3 = 0$, that is equivalent to state that the x-axis of the second camera is parallel to the baseline $\square$.

We can now demonstrate *Theorem 1*.

Proof (Theorem 1): by *Lemma 1* and *Lemma 2*, if the epipolar lines of the second camera are horizontal, the fundamental matrix can be reduced to the form:

$$\overline{\mathbf{F}} \propto \begin{bmatrix} \mathbf{0} \\ -[\mathcal{H}]_{3*} \\ [\mathcal{H}]_{2*} \end{bmatrix} \tag{24}$$

Using Eq. (3) we impose:

$$\begin{aligned} [\mathcal{H}]_{2*} &= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \\ [\mathcal{H}]_{3*} &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{25}$$

That is equivalent to require that the y and z axes of the first CCS, as seen from the second CCS system of reference (respectively $[\mathcal{H}]_{2*}$ and $[\mathcal{H}]_{3*}$), must be parallel to the corresponding axes of the second camera $\square$.

3.2 Analytic Derivation of the Minimising Rectification

By Eqs. (14) and (15), finding the global minimum of Eq. (10) is equivalent to find the optimal common orientation $\mathbf{R}^{new}$ of the virtual cameras. Since $\hat{\mathbf{x}}$ is imposed by the baseline, the problem is reduced to finding the $\hat{\mathbf{z}}$ so that distortion is minimised. Setting $\hat{\mathbf{y}} = \hat{\mathbf{z}} \times \hat{\mathbf{x}}$, will then determine $\mathbf{R}^{new}$.

Choosing the minimising $\hat{\mathbf{z}}$, in turn, is equivalent, by construction, to finding the pair of corresponding epipolar lines to become the new *horizon lines* (see Fig. 2), which will lay at the intersection of the plane $\hat{\mathbf{y}} = 0$ and the rectified images. Actually only one *horizon line* must be determined, as the corresponding one is readily found by means of $\mathbf{F}$.

Finally, the minimisation problem is reduced to a single parameter problem, noticing that an epipolar line is uniquely determined by its y-intercept $(0, y_1) \in I_1$ in ICS. Such y-intercept, in WCS, takes the form:

$$\mathbf{y}_1 = \mathbf{R}_1^{-1} \cdot (\mathbf{A}_1^{-1} \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T - \mathbf{t}_1) \tag{26}$$

The $\hat{\mathbf{z}}$ axes of the virtual cameras will thus be the direction of the line $\mathbf{z}$ perpendicular to $\mathbf{b}$, passing through $\mathbf{y}_1$, calculated as the difference between the vector going from $\mathbf{o}_1$ to $\mathbf{y}_1$ and its projection on the baseline (its derivation is done by means of the outer product $\otimes$ [21]):

$$
\begin{aligned}
\mathbf{z} &= (\mathbf{y}_1 - \mathbf{o}_1) - [(\mathbf{y}_1 - \mathbf{o}_1)^T \cdot \hat{\mathbf{x}}]\hat{\mathbf{x}} \\
&= (\mathbf{y}_1 - \mathbf{o}_1) - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}} \cdot (\mathbf{y}_1 - \mathbf{o}_1)
\end{aligned}
\tag{27}
$$

Then follows:

$$
\hat{\mathbf{z}} = \frac{\mathbf{z}}{\|\mathbf{z}\|}
\tag{28}
$$

Using Eqs. (14), (15) and (27), yields:

$$
\begin{aligned}
\mathbf{w}_1^T &= [\mathbf{H}_1]_{3*} \\
&= [\mathbf{K}_1 \cdot \mathbf{R}^{new} \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1}]_{3*} \\
&= \hat{\mathbf{z}}^T \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \\
&= \frac{\mathbf{z}^T}{\|\mathbf{z}\|} \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \\
&= \left(\mathbf{R}_1^{-1} \cdot \mathbf{A}_1^{-1} \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix} - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}} \cdot \mathbf{R}_1^{-1} \cdot \mathbf{A}_1^{-1} \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix} \right)^T \\
&\quad \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \\
&= \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-T} \cdot (\mathbb{I}_3 - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}}) \\
&\quad \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1}
\end{aligned}
\tag{29}
$$

where $\mathbf{K}_1$ has been omitted as it does not affect the 3rd row of what follows (being an affine transformation, its last row is $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$), and $\|\mathbf{z}\|$ has been discarded noticing that in Eq. (11) it appears both at numerator and denominator. Rearranging:

$$
\begin{aligned}
\mathbf{w}_1 &= (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-T} \cdot (\mathbb{I}_3 - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}}) \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1} \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix} \\
&= \mathbf{L}_1 \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix}
\end{aligned}
\tag{30}
$$

where the matrix:

$$
\mathbf{L}_1 = (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-T} \cdot (\mathbb{I}_3 - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}}) \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1}
\tag{31}
$$

Similarly for $\mathbf{w}_2$:

$$
\mathbf{w}_2 = \mathbf{L}_2 \cdot \begin{bmatrix} 0 \\ y_1 \\ 1 \end{bmatrix}
\tag{32}
$$

where:

$$
\mathbf{L}_2 = (\mathbf{A}_2 \cdot \mathbf{R}_2)^{-T} \cdot (\mathbb{I}_3 - \hat{\mathbf{x}} \otimes \hat{\mathbf{x}}) \cdot (\mathbf{A}_1 \cdot \mathbf{R}_1)^{-1}
\tag{33}
$$

Our goal is to minimise the distortion from Eq. (11), that can now be written as:

$$\frac{\begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot \mathbf{M}_1 \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T}{\begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot \mathbf{C}_1 \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T} + \frac{\begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot \mathbf{M}_2 \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T}{\begin{bmatrix} 0 & y_1 & 1 \end{bmatrix} \cdot \mathbf{C}_2 \cdot \begin{bmatrix} 0 & y_1 & 1 \end{bmatrix}^T} \tag{34}$$

where:

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{L}_1^T \cdot \mathbf{P}_1 \cdot \mathbf{P}_1^T \cdot \mathbf{L}_1 \\ \mathbf{M}_2 &= \mathbf{L}_2^T \cdot \mathbf{P}_2 \cdot \mathbf{P}_2^T \cdot \mathbf{L}_2 \\ \mathbf{C}_1 &= \mathbf{L}_1^T \cdot \mathbf{P}_{c1} \cdot \mathbf{P}_{c1}^T \cdot \mathbf{L}_1 \\ \mathbf{C}_2 &= \mathbf{L}_2^T \cdot \mathbf{P}_{c2} \cdot \mathbf{P}_{c2}^T \cdot \mathbf{L}_2 \end{aligned} \tag{35}$$

The terms of Eq. (34) can be now expanded as polynomial expressions in y_1. It can be verified that Eq. (34) takes the form:
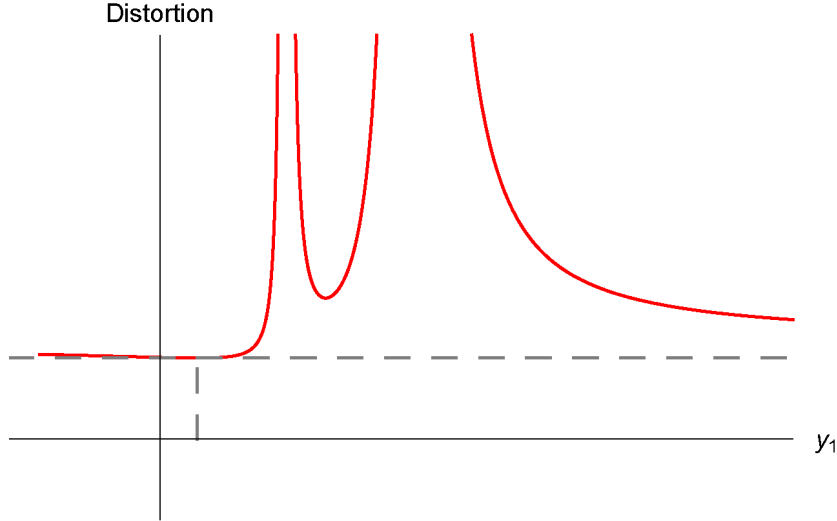

Figure 3: Possible trend of the distortion in Eq. (34) as a function of y_1. Global minimum is identified by dashed lines.

$$\frac{f_1(y_1)}{(g_1(y_1))^2} + \frac{f_2(y_1)}{(g_2(y_1))^2} \tag{36}$$

with $f_1 = f_1(y_1)$, $f_2 = f_2(y_1)$ second degree polynomials and $g_1 = g_1(y_1)$, $g_2 = g_2(y_1)$ first degree polynomials. Deriving to find the extreme points yields:

$$\left(\frac{f_1}{g_1^2} + \frac{f_2}{g_2^2} \right)' = \frac{g_1 g_2^3 f_1' + g_1^3 g_2 f_2' - 2 f_1 g_2^3 g_1' - 2 f_2 g_1^3 g_2'}{g_1^3 g_2^3} \tag{37}$$

11

where the terms of 5th degree cancel out. Discarding the denominator[1], the extreme points are thus found as the solution of a 4th degree polynomial:

$$ay_1^4 + by_1^3 + cy_1^2 + dy_1 + e = 0 \tag{38}$$

with $a, b, c, d, e \in \mathbb{R}$. The solutions of Eq. (38) can be found using any solving formula for homogeneous polynomials of 4th degree. Full calculations are reported in Appendix A.

Fig. 3 shows a possible behavior of Eq. (34). Among the acceptable real solutions of Eq. (38), the one representing the global minimum depends on the initial position of the cameras, and therefore can only be determined by directly comparing the value of the distortion in Eq. (34) for each solution.

It must be remarked that, in the very peculiar (unrealistic) case in which the two cameras are identical (i.e. $\mathbf{A}_1 = \mathbf{A}_2$, $\mathbf{P}_1 = \mathbf{P}_2$ and $\mathbf{P}_{c1} = \mathbf{P}_{c2}$) and share the exact same orientation $\mathbf{R}_1 = \mathbf{R}_2$ (not necessarily frontoparallel), then (36) takes the form:

$$2 \frac{f(y_1)}{(g(y_1))^2} \tag{39}$$

leading, once derived, to a first degree polynomial, thus to a single minimum.

The analytical expression of $\mathbf{w}_1$ and $\mathbf{w}_2$ minimising perspective distortion is therefore found, so that $\mathbf{H}_{1p}$ and $\mathbf{H}_{2p}$ are directly determined. In order to find the complete expression of the rectifying homographies $\mathbf{H}_1$ and $\mathbf{H}_2$, the respective affine components, as shearing and similarity transformations, can be easily calculated [1].

4 Algorithm Summary

The steps of the direct rectifying algorithm explained above are summarised as follows:

1. The weights $\mathbf{w}_1$ and $\mathbf{w}_2$ of the perspective components of the rectifying homographies are written as polynomial expression in y_1, the y-intercept of the *horizon* epipolar line on I_1.

2. Distortion is written as function of y_1 and auxiliary matrices are defined to calculate the coefficients of the quartic polynomial.

3. The acceptable solutions (either 2 or 4) are calculated.

4. The global minimum is determined by direct comparison of the distortion values obtained.

5. The minimising $\mathbf{w}_1$ and $\mathbf{w}_2$ are calculated as last row of the matrices in Eq. (15).

6. The similarity and shearing transformations are calculated [1].

7. The rectification transformations $\mathbf{H}_1$ and $\mathbf{H}_2$ are fully determined.

The full Python 3 code is made available at `https://github.com/decaden za/DirectStereoRectification`.

---

[1]It can be shown that approaching the roots of the denominator of the distortion function both from the upper and lower limit, the function always goes to $+\infty$, therefore the global minimum never reaches $-\infty$, as in Fig. 3.

5   Discussion

Rectifying homographies are transformations that dramatically reduce the search space for correspondences between a couple of stereo images, thus making the stereo matching problem computationally affordable.

The optimal transformations are traditionally found making use of numerical minimisation (see [1]), which prevents the use of rectification when low computational power is available (e.g. space assets). Furthermore, the application of the traditional method is limited by a convergence issue, due to the matrices decomposition used therein and thus not avoidable, that will be discussed in the reminder of this Section. For both these reasons, most present applications [22, 23] prefer to use the arbitrary, non-optimal solution proposed by Fusiello et al. [6], that, for peculiar configurations introduces high distortion levels (Sec. 5.1), thus impairing performance of the following stereo matching algorithm.

This paper explicits and demonstrates the formula for the optimal rectifying homograpies. The formula is valid for every pair of stereo images, independently from the configuration parameters. Being an exact formulation, it eliminates the need for minimisation, while still providing the optimal transformation, thus enabling the use of rectification in scenarios with very limited computational capabilities, as for example for autonomous navigation of space satellites. Future work will concentrate on such applications.

The reminder of this section discusses a limiting convergence issue of the traditional method proposed by Loop and Zhang [1]. Indeed, the convergence of this algorithm cannot be guaranteed as it depends upon finding a suitable initial guess, derived assuming $\mathbf{A}$ and $\mathbf{A}'$ to be positive-definite, which cannot be guarenteed for all configurations [7].

Consider, for example, the case in which the first CCS is coincident with the WCS and an identical second camera is placed in $\begin{bmatrix} 1 & a & b \end{bmatrix}^T$ (with $a, b \in \mathbb{R}$), oriented as the first one but rotated around the x-axis of an angle θ_x. For all cases in which $b = a \tan \theta_x$, the first element of $\mathbf{A}'$ will be null, thus, as a straightforward consequence of *Sylvester's Theorem*, $\mathbf{A}'$ will not be positive-definite. Then, Loop and Zhang's algorithm fails.

Moreover, numerical simulations confirm that the configurations for which positive-definite assumption is violated are a numerous and therefore relevant in limiting the applicability of the traditional algorithm. Indeed, considering a setup with fixed intrinsic and randomly generating one million of extrinsic parameters (i.e. relative position and rotation between the two cameras), it was found that Loop and Zhang's algorithm failed in over the 50% of the cases, in spite of all the configurations being perfectly legitimate. This happened mostly because of numerical errors in computing Cholesky decomposition of $\mathbf{A}$ and $\mathbf{A}'$. The proposed analytic method, instead, does not need an initial guess and directly provides the optimal solution.

To the best of the authors's knowledge, for the cases mentioned above, all algorithms present in the the literature fail in providing the optimal rectifying homography (introducing minimum distortion), found instead by the formula derived in this work and not affected by the configuration settings.

Furthermore, Fig. 4 shows the optimal solution found by our algorithm in case of extremely skewed camera positions configuration, for which Loop and Zhang's initial guess cannot be retrieved. Finally it must be mentioned that

the direct algorithm here presented is capable of providing the optimal solution also in the case of one camera entering the field of view of the other, i.e. when the epipoles lie within image boundaries, which causes convergence issues when employing numerical algorithms.
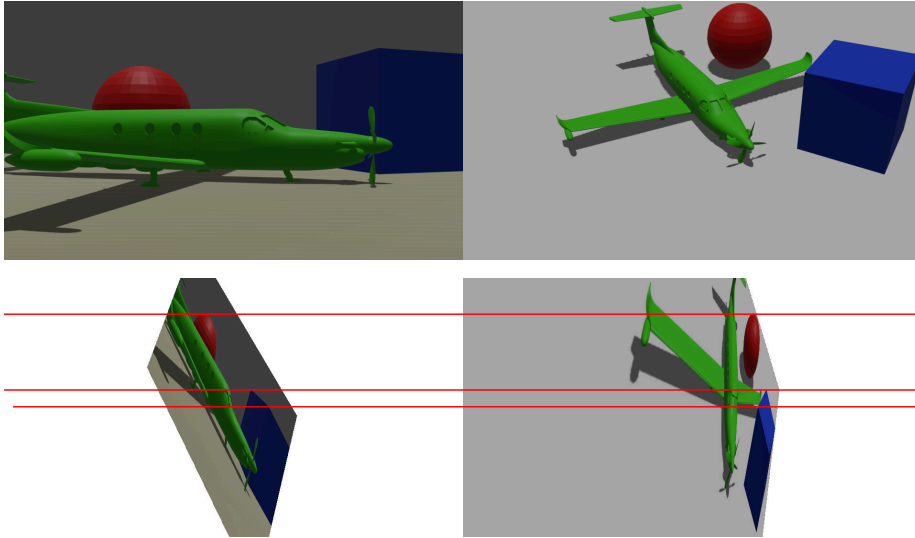

Figure 4: Left and right image of a synthetic scene with extremely skewed camera positions (top-left and top-right, respectively). In this case the numerical minimisation of [1] fails. Corresponding rectified image pair using our algorithm is showed (bottom). Horizontal lines are drawn for reference on the rectified image.

In summary, the formula here presented guarantees convergence for every configuration, avoids the need of external libraries and saves computational time as it does not make use of minimisation. Indeed, comparing the computational cost, in our implementation on an Intel i7-7500U CPU 2.70GHz, the original algorithm by Loop and Zhang requires 8.899 ms on average to calculate the rectifying homographies, while the proposed method takes 7.655 ms.

5.1 Rectification Examples

A synthetic example of rectification is shown in Fig. 5. The original left and right images are first shown. The images, as rectified following our direct algorithm, are then listed, where arbitrary horizontal lines have been drawn as a reference. Here the effects of rectification are clearly visible, as corresponding points are aligned. This example is included in the Python code.

The couple of images on the third line of Fig. 5, shows the same image pair rectified following the algorithm in [6]. As expected, while the value of the minimum distortion introduced in the second row is 46 252, the homographies found by Fusiello's algorithm and generating the third pair of images are non optimal and introduce a distortion of 48 207, about 4% higher than the minimum.

In Fig. 6 a real scene stereo pair is rectified. Despite the cameras are placed almost frontoparallel, corresponding points still lie on different scan-lines, so
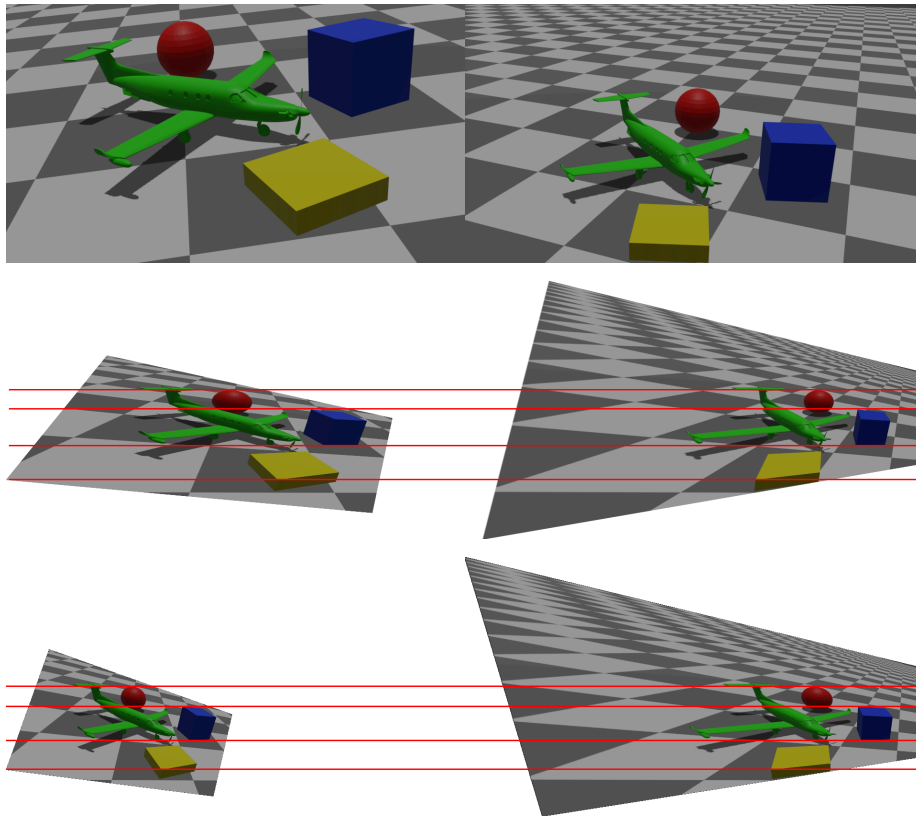
Figure 5: Left and right image of a synthetic scene (top-left and top-right, respectively) and corresponding rectified image pair as rectified using the direct algorithm (center) and the algorithm in [6](bottom). Horizontal lines are drawn for reference on the rectified images.

rectification is needed. In real applications, calibration [2] is required to accurately fit the stereo rig to the pinhole camera model. Apart from lens distortion correction, the algorithm finds no difference between real-case and synthetic images, since it starts from the same premises (i.e. both camera projection matrices).

6   Conclusions

A direct rectification algorithm for calibrated stereo rigs has been proposed. Our method improves the well-known approach by Loop and Zhang to calculate the optimal rectifying homographies.

Thanks to an alternative geometrical interpretation of the problem, the proposed algorithm is able to find the formula for the rectifying homographies that introduce the minimal perspective distortion for any camera configurations, even in extremely skewed relative positions, and without depending on minimisation libraries.

The Python 3 code has been made publicly available at `https://github.c`

Figure 6: Left and right image of a real scene (top-left and top-right, respectively) and corresponding rectified image pair (bottom). Horizontal lines are drawn for reference on both images, showing effects of rectification. Lens distortion correction is also noticeable.

om/decadenza/DirectStereoRectification.

Because of the lower computational cost, the value of having an analytic solution may be particularly relevant for hardware-limited applications (e.g. space applications, miniaturised devices), where each change of camera extrinsic or intrinsic parameters would require the computation of new rectifying homographies. Future work might formulate the analytic solution for a distortion metric that includes pixel resampling effects and applications to scenarios with very limited computational capabilities.

References

[1] C. Loop and Z. Zhang, "Computing rectifying homographies for stereo vision," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, vol. 1, 1999.

[2] Z. Zhang, "A flexible new technique for camera calibration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2000.

[3] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision.* Cambridge, UK: Cambridge University Press, 2 ed., 2003.

[4] Hornberg, *Handbook of Machine and Computer Vision.* Wiley-VCH, 2017.

[5] P. Besl, *Active Optical Range Imaging Sensors.* New York, NY: Springer, 1989.

[6] A. Fusiello, E. Trucco, and A. Verri, "A compact algorithm for rectification of stereo pairs," *Machine Vision and Applications*, 2000.

[7] L. Sui, Z. Jiulong, and C. Duwu, "Image rectification using affine epipolar geometric constraint," 2009.

[8] J. Gluckman and S. Nayar, "Rectifying transformations that minimize resampling effects," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, 2001.

[9] C. Sun, "Closed-form stereo image rectification," Association for Computing Machinery, 2012.

[10] Z. Chen, C. Wu, and H. Tsui, "A new image rectification algorithm," *Pattern Recognition Letters*, 01 2003.

[11] S. Kumar, C. Micheloni, C. Piciarelli, and G. Foresti, "Stereo rectification of uncalibrated and heterogeneous images," *Pattern Recognition Letters*, 2010.

[12] X. Ruichao, S. Wenxiu, P. Jiahao, Y. Qiong, and R. Jimmy, "Dsr: Direct self-rectification for uncalibrated dual-lens cameras," *International Conference on 3D Vision*, 2018.

[13] R. Deriche, Z. Zhang, Q.-T. Luong, and O. Faugeras, "Robust recovery of the epipolar geometry for an uncalibrated stereo rig," *Proceedings of the Third European Conference on Computer Vision, vol. 1, Stockholm, Sweden*, 2001.

[14] P. Monasse, J.-M. Morel, and Z. Tang, "Three-step image rectification," in *BMVC 2010-British Machine Vision Conference*, 2010.

[15] S. Abraham and W. Förstner, "Fish-eye-stereo calibration and epipolar rectification," *Isprs Journal of Photogrammetry and Remote Sensing*, 2005.

[16] X. Zhucun, N. Xue, G.-S. Xia, and W. Shen, "Learning to calibrate straight lines for fisheye image rectification," 04 2019.

[17] X. Yin, X. Wang, J. Yu, M. Zhang, P. Fua, and D. Tao, "Fisheyerecnet: A multi-context collaborative deep network for fisheye image rectification," in *Proceedings of the European conference on computer vision (ECCV)*, 2018.

[18] Q.-T. Luong and O. Faugeras, "The fundamental matrix: Theory, algorithms, and stability analysis," *International Journal of Computer Vision*, 1996.

[19] J. Gallier, *Basics of Projective Geometry*. Springer New York, 2011.

[20] C. Turrini, "Geometria per la ricostruzione tridimensionale," 2017.

[21] H. Lu, K. Plataniotis, and A. Venetsanopoulos, *Multilinear Subspace Learning: Dimensionality Reduction of Multidimensional Data*. Chapman & Hall/CRC, 01 2013.

[22] R. Xia, R. Su, J. Zhao, Y. Chen, S. Fu, L. Tao, and Z. Xia, "An accurate and robust method for the measurement of circular holes based on binocular vision," *Measurement Science and Technology*, 2019.

[23] T. Hinzmann, J. L. Schönberger, M. Pollefeys, and R. Siegwart, "Mapping on the fly: Real-time 3d dense reconstruction, digital surface map and incremental orthomosaic generation for unmanned aerial vehicles," in *Field and Service Robotics*, 2018.

Appendix A

The coefficients of the 4th order polynomial expression in y_1 of Eq. (38) are:

$$
\begin{aligned}
a &= m_2 m_4 + m_6 m_8 \\
b &= m_1 m_4 + 3 m_2 m_3 m_4 + m_5 m_8 + 3 m_6 m_7 m_8 \\
c &= 3 m_1 m_3 m_4 + 3 m_2 m_3^2 m_4 + 3 m_5 m_7 m_8 + 3 m_6 m_7^2 m_8 \\
d &= 3 m_1 m_3^2 m_4 + m_2 m_3^3 m_4 + 3 m_5 m_7^2 m_8 + m_6 m_7^3 m_8 \\
e &= m_1 m_3^3 m_4 + m_5 m_7^3 m_8
\end{aligned}
\tag{40}
$$

with:

$$
\begin{aligned}
m_1 &= [\mathbf{M}_1]_{23} [\mathbf{C}_1]_{23} - [\mathbf{M}_1]_{33} [\mathbf{C}_1]_{22} \\
m_2 &= [\mathbf{M}_1]_{22} [\mathbf{C}_1]_{23} - [\mathbf{M}_1]_{23} [\mathbf{C}_1]_{22} \\
m_3 &= \frac{[\mathbf{C}_2]_{23}}{[\mathbf{C}_2]_{22}} \\
m_4 &= \frac{[\mathbf{C}_2]_{22}}{[\mathbf{C}_1]_{22}} \\
m_5 &= [\mathbf{M}_2]_{23} [\mathbf{C}_2]_{23} - [\mathbf{M}_2]_{33} [\mathbf{C}_2]_{22} \\
m_6 &= [\mathbf{M}_2]_{22} [\mathbf{C}_2]_{23} - [\mathbf{M}_2]_{23} [\mathbf{C}_2]_{22} \\
m_7 &= \frac{[\mathbf{C}_1]_{23}}{[\mathbf{C}_1]_{22}} \\
m_8 &= \frac{[\mathbf{C}_1]_{22}}{[\mathbf{C}_2]_{22}}
\end{aligned}
\tag{41}
$$

The four roots of the equation are given by:

$$
y_1 = \frac{-b}{4a} \pm Q \pm \frac{1}{2} \sqrt{-4Q^2 - 2p + \frac{S}{Q}}
\tag{42}
$$

with:

$$Q = \frac{1}{2}\sqrt{-\frac{2}{3}p + \frac{1}{3a}\left(\Delta_0 + \frac{q}{\Delta_0}\right)}$$

$$S = \frac{8a^2d - 4abc + b^3}{8a^3}$$

$$\Delta_0 = \left(\frac{s + \sqrt{s^2 - 4q^3}}{2}\right)^{\frac{1}{3}} \tag{43}$$

$$p = \frac{8ac - 3b^2}{8a^2}$$

$$q = 12ae - 3bd + c^2$$

$$s = 27ad^2 - 72ace + 27b^2e - 9bcd + 2c^3$$

Remark: for the case $\mathbf{A}_1 = \mathbf{A}_2$, $\mathbf{P}_1 = \mathbf{P}_2$, $\mathbf{P}_{c1} = \mathbf{P}_{c2}$ and $\mathbf{R}_1 = \mathbf{R}_2$, the solution is given by:

$$y_1 = -\frac{m_1}{m_2} \tag{44}$$

2022-07-07

Rectifying homographies for stereo vision: analytical solution for minimal distortion

Lafiosca, Pasquale

Springer