We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

13

Projective Rectification with Minimal Geometric Distortion

Hsien-Huang P. Wu and Chih-Cheng Chen
National Yunlin University of Science and Technology
Taiwan

1. Introduction

There has been an increasing interest in the 3D imaging in the fields of entertainment, simulation, medicine, 3D visual communication, 3D tele-robotics, and 3D TV to augment the reality of presence or to provide vivid and accurate structure information. In order to provide vivid information in these and other 3D applications, efficient techniques to generate, store, and view the stereoscopic video are essential. While many methods are available for acquiring stereoscopic video, the images pairs obtained might not be in rectified form. Therefore, rectification is usually needed to support comfortable viewing and effective compression for storage and transmission. Projective geometry has been proved to be a useful tool for solving the rectification problem without camera calibration. However, if the matrices used for projective rectification (homographies) are not constrained properly, the rectification process can cause great geometric distortion. For visual applications, rectification with minimum geometry distortion should be pursued. In this chapter, we propose an improved algorithm to minimize the distortion by combining a newly developed projective transform with a properly chosen shearing transform. This new method is equipped with flexibility and can be adapted to various imaging models. Experimental data show that the proposed method works quite well for all the image pairs taken different imaging conditions. Comparison with other available method based on visual inspection and numerical data demonstrates the superiority of the new approach.

2. Background

Stereo vision is a technique for estimating 3D structure based on two or more images taken from different viewpoints, and are most often used in robotics and vehicle navigation. Stereoscopic videos are vastly used in entertainment, gaming, simulation, tele-conferencing, and tele-operation to augment the reality of presence. One of the major issues in the application of stereo imagery is correspondence problem. The correspondence problem is defined as locating a pair of image pixels from two different images, where these two pixels are projections of the same scene element. Given a point in one image, its *correspondent point* (or point correspondence) must lie on an epipolar line in the other image. This relationship is well known as epipolar constraint (Faugeras, 1993). It is obvious that knowledge of this epipolar geometry, or codified as *fundamental matrix*, simplifies the stereo matching from a 2-D area search to a 1-D search along the epipolar line. If the images are acquired from a

Source: Scene Reconstruction, Pose Estimation and Tracking, Book edited by: Rustam Stolkin,
ISBN 978-3-902613-06-6, pp.530, I-Tech, Vienna, Austria, June 2007

pair of identical cameras placed side-by-side and pointed in the same direction, known as a *rectilinear stereo rig*, the epipolar lines will coincide with scan lines (x-axis) of the images. Given this ideal epipolar geometry, the correspondent points will lie on the same scan line in the two images. However, for an arbitrary placement of cameras, the epipolar lines are skew and the 1-D search will still be time consuming. Whether the imagery is used for stereo vision or stereoscopic video, we would like the image pair to be taken from an ideal epipolar-geometry.

When the epipolar geometry is not in ideal form, the image pairs can be warped to make correspondent points lie on the same scan lines. This process is known as image rectification, and can be accomplished by applying a 2D projective transforms, or *homographies*, to each image. The homography is a linear one to one transformation of the projective plane, which is represented by a 3×3 non-singular matrix. The rectified images can then be treated as obtained by a rectilinear stereo rig and the correspondence problem is greatly simplified. Since most stereo algorithms assume input images having ideal epipolar geometry, image rectification is usually a pre-requisite operation for stereoscopic related applications.

The idea of rectification has long been used in photogrammetry (Slama, 1980). The techniques originally used were optical-based, but now are replaced by software methods that model the geometry of optical projection. The software-based photogrammetric approaches, similar to most of the computer vision ones, assume the knowledge of projection matrices or cameras parameters (Ayache & Hanse, 1988)(Ayache & Lustman, 1991)(Fusiello, et al., 2000)These methods require camera parameters to compute a pair of homographies for transformations. The necessity of camera calibration is one of their disadvantages.

In contrast to these traditional approaches, several researchers have developed techniques called *projective rectification* to rectify images directly without using camera parameters. They utilized the epipolar geometry of the acquired images and various criteria to compute the homographies. Robert *et al.* (Robert, 1997) attempted to find the transform that best preserves orthogonality around image centers. Hartley (Hartley, 1999) proposed using minimization of the differences between matching points for the solution of homographies. He also gave a detailed theoretical presentation of the projective rectification. Loop and Zhang (Loop & Zhang, 1999) suggested decomposing each homography into projective and affine components. They then found the projective component that minimizes a defined projective distortion criterion. Gluckman and Nayar (Gluckman & Nayar, 2001) recently presented a stereo rectification method, which takes geometric distortion into account and tries to minimize the effects of resampling. Pollefeys (Pollefeys et al., 1999) proposed a simple and efficient algorithm for general two view stereo image. The other available approaches include (Papadimitriou and Dennis, 1996) which considers only the special case of partially aligned cameras, and (Al-Shalfan et al., 2000) which requires the estimation of the epipolar geometry. Although these proposed methods provided many possibilities for projective rectification, they all solve the problem indirectly. That is, they must explicitly estimate the fundamental matrix before rectification. Since the solution of fundamental matrix has its own uncertainty (Zhang, 1998) this indirect approach might obtain unpredictable rectifying results.

Isgrò and Trucco (Isgrò & Trucco, 1999) adopted a different procedure and obtained homographies directly without first computing the fundamental matrix. However, in order

to solve the problem of uniqueness in rectification, their method requires disparity minimization along the *x-axis* to generate a unique solution. In certain applications, this modification of *x-axis* disparity in rectification might be harmless; however, in applications where original *x-axis* disparity must be maintained (e.g. for stereoscopic viewing purpose), this constraint will make the algorithm useless. Moreover, the enforcement of minimizing *x-axis* disparity to obtain a single solution sometimes greatly distorts the image.

In this chapter, we propose a different approach for rectifying two uncalibrated images with reduced geometric distortion. Its novelty is to formulate a new set of parameters for homographies, and solves the rectification problem using least square distance as a criterion. This new approach possesses similar advantage to that of the IT method (Isgrò & Trucco, 1999), that is, performing uncalibrated rectification without explicit computation of the epipolar geometry (fundamental matrix). However, the new method contains a shearing transform which greatly reduces geometric distortion caused by rectification.

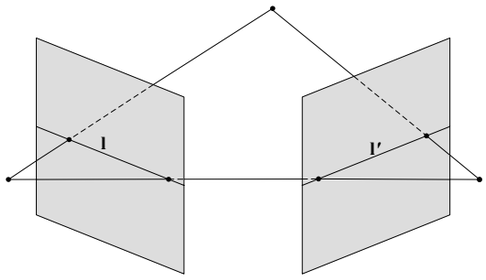


Figure 1. Epipolar geometry of a pair of stereo images

3. Epipolar geometry

The derivation of our algorithm is presented from the viewpoint of projective geometry (Faugeras, 1993). The image point is expressed in homogeneous coordinate and represented by a 3-dimensional column vector. Column vectors are denoted by bold lower-case letters, such as $\mathbf{m}$ and $\mathbf{l}$. Matrices are represented by bold upper-case letter, such as $\mathbf{F}$ and $\mathbf{H}$. Transposed vectors and matrices are expressed by adding a letter T as superscript, e.g., $\mathbf{m}^T$ and $\mathbf{F}^T$.

3.1 Epipolar constraint

Consider two images I and I' of a common scene. Let C and C' represent the optical centers of the left and right cameras in the 3D coordinate, respectively. Points $\mathbf{m}$ and $\mathbf{m}'$ are the projections of a certain 3D point M on the left (I) and right (I') images, as shown in Fig. 1. There are two points called the epipoles of the left and right images; one epipole $\mathbf{e}'$ is the point where the center of projection C of the left camera would be visible in the right image, and the other epipole $\mathbf{e}$ is the point where the center of projection C' of the right camera is seen in the left image. In the 3D coordinate, $\mathbf{e}$ and $\mathbf{e}'$ are the intersection points of

the baseline, $\overline{CC'}$, with the left and right image planes. Any plane that contains the baseline and a 3D scene point (e.g. M above) is called *epipolar plane*.

The *epipolar lines* ($\mathbf{l}$ and $\mathbf{l}'$ in Fig. 1) are defined as the intersection of the epipolar plane and the left and right image planes. The ray goes through the optical center of one camera (C), which creates an image point ($\mathbf{m}$) on the image plane, will generate an image of epipolar line on the other camera ($\mathbf{l}'$). One specific matrix called *fundamental matrix* describes this mapping between points in one image and the corresponding epipolar line in the other image. Given the scene point *M*, and its two projections ($\mathbf{m}$ and $\mathbf{m}'$) on the left and right image planes, the *epipolar constraint* (Faugeras, 1993) asserts that point $\mathbf{m}'$ must lie on the epipolar line $\mathbf{Fm}$ and can be expressed as

$$\mathbf{m}'^T\mathbf{F}\mathbf{m} = \mathbf{0} \quad \text{or} \quad \mathbf{m}^T\mathbf{F}^T\mathbf{m}' = \mathbf{0} \tag{1}$$

Where $\mathbf{F}$ is called fundamental matrix, or *F matrix*, and $\mathbf{0} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$ is a zero column vector. The matrix $\mathbf{F}$ is a 3×3 matrix with rank 2, and the epipoles for the left ($\mathbf{e} \in I$) and the right image ($\mathbf{e}' \in I'$) satisfy

$$\mathbf{F}\mathbf{e} = \mathbf{F}^T\mathbf{e}' = \mathbf{0} \tag{2}$$

That is, the epipoles $\mathbf{e}$ and $\mathbf{e}'$ are the null space of $\mathbf{F}$ and $\mathbf{F}^T$, respectively. Furthermore, all the epipolar lines ($\mathbf{l}$ and $\mathbf{l}'$) will pass the epipoles, or

$$\mathbf{l}^T\mathbf{e} = \mathbf{l}'^T\mathbf{e}' = 0 \tag{3}$$

3.2 Epipolar geometry after stereo image rectification

Image rectification can be treated as a process of converting the epipolar geometry of an image pair into a canonical form. This can be done by applying a homography, which maps the epipole to a point at infinity, to each image. We designate these two epipoles after rectification as $\mathbf{e}_\infty$ and $\mathbf{e}'_\infty$, where $\mathbf{e}_\infty = \mathbf{e}'_\infty = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$, and fundamental matrix of a pair of rectified images has the form of

$$\mathbf{F}_\infty = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \tag{4}$$

Let $(\bar{\mathbf{m}}, \bar{\mathbf{m}}')$ be an image pair in the rectified images corresponding to the original $(\mathbf{m}, \mathbf{m}')$ pair. From equation (1), the epipolar constraint after rectification can be rewritten as

$$\bar{\mathbf{m}}'^T\mathbf{F}_\infty\bar{\mathbf{m}} = 0 \tag{5}$$

Meanwhile, the rectification can be accomplished by the following operations:

$$\bar{\mathbf{m}} = \mathbf{H}\mathbf{m}\,, \quad \bar{\mathbf{m}}' = \mathbf{H}'\mathbf{m}' \tag{6}$$

where $\mathbf{H}$ and $\mathbf{H}'$ are the rectifying *homographies* (or *H matrices*) for the left and right images, respectively. Combining (5) and (6) we obtain the following equation:

$$\mathbf{m}'^{T}\mathbf{H}'^{T}\mathbf{F}_{\infty}\mathbf{H}\mathbf{m} = 0 \tag{7}$$

Given several point correspondences, or $(\mathbf{m}, \mathbf{m}')$ pairs, equation (7) can be solved to obtain homographies ($\mathbf{H}$ and $\mathbf{H}'$). However, solution for the pair of *H matrices* $(\mathbf{H}, \mathbf{H}')$ is not unique. Some of the solutions are even far from ideal and can cause huge geometric distortion. Various approaches have been proposed to find a unique pair of homographies $(\mathbf{H}, \mathbf{H}')$ that minimize image distortion. This and other rectification related backgrounds are the topics of the next section.

4. Image Rectification

Several projective rectification methods have been proposed recently, and the backgrounds on these methods that are most related to our algorithm will be described below. On the basis of these methods, we propose our new approach to solving the projective rectification problem.

4.1 Review of projective rectification

Rectification based on epipolar geometry was originally developed by Hartley (Hartley, 1999). In order to constrain the geometric distortion caused by rectification, Hartley proposed that one of the two homographies, say $\mathbf{H}'$, should be close to a rigid transformation in the neighborhood of a selected point $\mathbf{p}_0$. That is, the homography for one of the image (I') can be represented by

$$\mathbf{H}' = \mathbf{KRT} \tag{8}$$

where $\mathbf{T}$ is a translational vector taking $\mathbf{p}_0$ to the origin, $\mathbf{R}$ is a rotation matrix mapping the epipole to a point $\begin{bmatrix} 1 & 0 & f \end{bmatrix}^T$ on the *x-axis,* and $\mathbf{K}$ is a transformation matrix mapping $\begin{bmatrix} 1 & 0 & f \end{bmatrix}^T$ to a point $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$ at infinity along the *x-axis* . Moreover, matrix $\mathbf{K}$ can be expressed as

$$\mathbf{K} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -f & 0 & 1 \end{bmatrix} \tag{9}$$

In this way, $\mathbf{H}'$ depends only on two parameters: f and rotation angle θ. If the translational vector $\mathbf{T}$ is neglected, then $\mathbf{H}'$ becomes

$$\mathbf{H}' = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ -f\cos\theta & -f\sin\theta & 1 \end{bmatrix} \tag{10}$$

Given the **F** matrix, positions of the epipoles can be found by Equation (2). By following the above process, we can then obtain $\mathbf{H}'$ to rectify the image I' by mapping the epipole $\mathbf{e}'$ to the infinite point $\begin{bmatrix}1 & 0 & 0\end{bmatrix}^T$ and transforming the epipolar lines to lines parallel with the x-axis. The next step is to find the matrix $\mathbf{H}$ which can be applied to the other image to match up these new epipolar lines.

The strategy that Hartley took to find $\mathbf{H}$ (a matching transformation) is to minimize the sum-of-squared distances (Hartley, 1999):

$$\sum_i d(\mathbf{H}\mathbf{m}_i, \mathbf{H}'\mathbf{m}_i')^2 \tag{11}$$

The searching of $\mathbf{H}$ by minimizing $\sum_i d(\mathbf{H}\mathbf{m}_i, \mathbf{H}'\mathbf{m}_i')^2$ is not as straightforward as it seems. That is, the matrix $\mathbf{H}$ is first decomposed into a form of $\mathbf{H} = \mathbf{H}_A\mathbf{H}_0$ where

$$\mathbf{H}_A = \begin{bmatrix} a & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } \mathbf{H}_0 = \mathbf{H}'\mathbf{M} \tag{12}$$

and $\mathbf{H}_A$ matrix represents an affine transformation. It was proven (Hartley, 1999) that *F matrix* can be factorized as $\mathbf{F} = [\mathbf{e}]_\times\mathbf{M}$ where $\mathbf{M}$ is a three-parameter family of non-singular matrices, and $[\mathbf{e}]_\times$ is an antisymetric matrix generated from vector $\mathbf{e}$ as follows:

$$[\mathbf{e}]_\times = \begin{bmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & -e_1 \\ -e_2 & e_1 & 0 \end{bmatrix} \tag{13}$$

Note that $\mathbf{e} = [e_1 \quad e_2 \quad e_3]^T$ is the epipole of the image I. Therefore, instead of minimizing $\sum_i d(\mathbf{H}\mathbf{m}_i, \mathbf{H}'\mathbf{m}_i')^2$ directly, the following steps are taken:

1. Matrix $\mathbf{H}'$ is found first.
2. The feature points on both images are transformed by

$$\begin{aligned} \hat{\mathbf{m}}_i &= \mathbf{H}_0\mathbf{m}_i = \mathbf{H}'\mathbf{M}\mathbf{m}_i \\ \hat{\mathbf{m}}_i' &= \mathbf{H}'\mathbf{m}_i' \end{aligned} \tag{14}$$

3. $\sum_i d(\mathbf{H}_A\hat{\mathbf{m}}_i, \hat{\mathbf{m}}_i')^2$ is minimized to find the matrix $\mathbf{H}_A$ by least square. (Note that this step would remove the x-disparity)
4. Matrix $\mathbf{H}$ is obtained by $\mathbf{H} = \mathbf{H}_A\mathbf{H}_0$

To further evaluate performance of the proposed algorithm, Hartley's method will be implemented and applied to the same sets of image pairs for comparisons in the later section of experiments. Algorithm of Hartley's approach has been briefly described above and is summarized below.

Outline of Hartley's Algorithm

1. Identify a set of point correspondences $\{\mathbf{m}_i \leftrightarrow \mathbf{m}'_i, i = 1 \ldots N\}$ between the two input images. Seven points at least are needed, though more are preferable.
2. Compute the fundamental matrix **F** and find the epipoles $\mathbf{e}$ and $\mathbf{e}'$ in the two images.
3. Select a projective transformation $\mathbf{H}'$ that maps the epipole $\mathbf{e}'$ to the point at infinity on the x-axis, $\begin{bmatrix}1 & 0 & 0\end{bmatrix}^T$.
4. Find the matching transformation H that minimizes the least-squares distance $\sum_i d(\mathbf{H}\mathbf{m}_i, \mathbf{H}'\mathbf{m}'_i)^2$.
5. Resample the first image according to the transformation $\mathbf{H}$ and the second image according to the projective transformation $\mathbf{H}'$.

The number of parameter needs to be estimated is ten in the above process which includes the computation of matrix **F** that requires estimation of seven parameters. The Matlab codes for implementation of the Hartley's method are available from (Hartley, 2004).

4.2 Proposed method and *F matrix* parameterization

In this section, taking a distinct approach in representing the matrix $\mathbf{H}$, we propose a new method to solving the two homographies. Unlike the way $\mathbf{H}$ is parameterized in as $\mathbf{H} = \mathbf{H}_A\mathbf{H}_0$ in Hartley's approach, the proposed method adopts a more direct form of parameterization for $\mathbf{H}$, that is

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \tag{15}$$

Since the homography pair, $\mathbf{H}$ and $\mathbf{H}'$, are determined up to a scale factor, we can set $\mathbf{H}(3,3) = \mathbf{H}'(3,3) = 1$. Combining equations (5), (6) and (7) gives us

$$\bar{\mathbf{m}}'^T\mathbf{F}_\infty\bar{\mathbf{m}} = \mathbf{m}'^T\mathbf{H}'^T\mathbf{F}_\infty\mathbf{H}\mathbf{m} = \mathbf{m}'^T\mathbf{F}\mathbf{m} = \mathbf{0}\text{, where } \mathbf{F} = \mathbf{H}'^T\mathbf{F}_\infty\mathbf{H} \tag{16}$$

By following Hartley's proposition that $\mathbf{H}'$ is very close to a rigid transformation, as shown in (10), and substituting $\mathbf{F}_\infty$, $\mathbf{H}'$, and $\mathbf{H}$ in equations (4), (10) and (15) into $\mathbf{F}$, we can parameterize and estimate *F matrix* as follows:

$$\hat{\mathbf{F}} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ -f\cos\theta & -f\sin\theta & 1 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -h_4 f\cos\theta + h_7\sin\theta & -h_5 f\cos\theta + h_8\sin\theta & -h_6 f\cos\theta + \sin\theta \\ -h_4 f\sin\theta - h_7\cos\theta & -h_5 f\sin\theta - h_8\cos\theta & -h_6 f\sin\theta - \cos\theta \\ h_4 & h_5 & h_6 \end{bmatrix} \tag{17}$$

Obviously, the *F matrix* is determined by only seven parameters, as shown in vector form

$$\mathbf{\Phi} = \begin{bmatrix} f & \theta & h_4 & h_5 & h_6 & h_7 & h_8 \end{bmatrix}^T \tag{18}$$

However, due to the characteristics of $\mathbf{F}_\infty$, only five out of eight parameters in matrix **H** are obtained, which leaves the solution for **H** not unique. To solve this uniqueness problem, Hartley's method suggests minimizing the discrepancy after rectification, as is described in Equation (11). In contrast, we propose using shearing transform formulated in section 3.4 to obtain a unique solution. Our approach results in lower geometric distortion as will be seen in the result section. Note that we have combined the problems of rectification and estimation of *F matrix*. In the next subsection, we will show how to derive a least-square solution for the rectification problem from the viewpoint of *F matrix* estimation. This novel parameterization scheme combining with a shearing transform, leads to a unique solution.

4.3 Projection rectification based on least square distance

The quantity used in Hartley's method for minimization, as shown in equation (11), is a linear criterion without physical meaning. In rectification, we would like the criterion to be something geometrically meaningful and to be measurable in the image plane. One such quantity is the distance from a point $\mathbf{m}_i'$ to its corresponding epipolar line $\mathbf{l}_i' = \mathbf{F}\mathbf{m}_i = \begin{bmatrix} l_1' & l_2' & l_3' \end{bmatrix}^T$, as shown in Fig. 2. This distance is given by the following equation:

$$d(\mathbf{m}_i', \mathbf{l}_i') = d(\mathbf{m}_i', \mathbf{F}\mathbf{m}_i) = \frac{\mathbf{m}_i'^T \mathbf{l}_i'}{\sqrt{l_1'^2 + l_2'^2}} = \frac{\mathbf{m}_i'^T \mathbf{F}\mathbf{m}_i}{\sqrt{l_1'^2 + l_2'^2}} \tag{19}$$

Conversely, distance for a point $\mathbf{m}_i$ to its corresponding epipolar line $\mathbf{l}_i = \mathbf{F}^T\mathbf{m}_i' = \begin{bmatrix} l_1 & l_2 & l_3 \end{bmatrix}^T$ is

$$d(\mathbf{m}_i, \mathbf{l}_i) = d(\mathbf{m}_i, \mathbf{F}^T\mathbf{m}_i') = \frac{\mathbf{m}_i^T \mathbf{l}_i}{\sqrt{l_1^2 + l_2^2}} = \frac{\mathbf{m}_i^T \mathbf{F}^T \mathbf{m}_i'}{\sqrt{l_1^2 + l_2^2}} \tag{20}$$

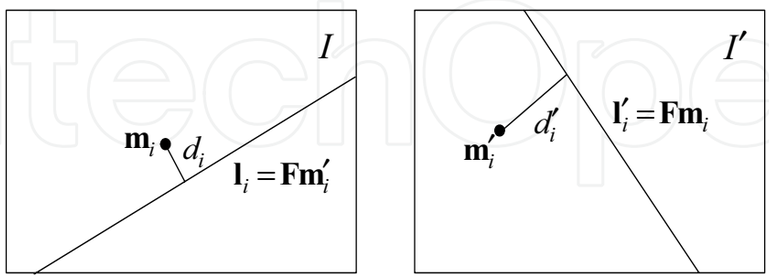


Figure 2. Distance from a point to the epipolar line of its correspondent point

Minimization of this distance is originally used to estimate *F matrix* (Zhang, 1998), which generates 36 possible modes of solution. Surprisingly, it turns out to be a very good *H matrix* estimator with a unique solution under our new formulation. We now demonstrate how to find the solution by minimizing the distance defined above.

To prevent inconsistency of the epipolar geometry between the left and right images, we choose to operate simultaneously on both images and minimize the mean-square distance. Hence, the problem becomes

$$\min_{\mathbf{F}} \sum_i \frac{1}{2}\left(d^2\left(\mathbf{m}_i', \mathbf{F}\mathbf{m}_i\right) + d^2\left(\mathbf{m}_i, \mathbf{F}^T\mathbf{m}'\right)\right) \tag{21}$$

Using equations (19), (20) and the fact that $\mathbf{m}_i'^T\mathbf{F}\mathbf{m}_i = \mathbf{m}_i^T\mathbf{F}^T\mathbf{m}_i'$, we reformulate the minimization problem in (21) as:

$$\min_{\mathbf{F}} \sum_i \frac{1}{2}\left[\frac{\left(\mathbf{m}_i'^T\mathbf{F}\mathbf{m}_i\right)^2}{l_1'^2 + l_2'^2} + \frac{\left(\mathbf{m}_i^T\mathbf{F}^T\mathbf{m}_i'\right)^2}{l_1^2 + l_2^2} \right] = \min_{\mathbf{F}} \sum_i w_i \frac{\left(\mathbf{m}_i'^T\mathbf{F}\mathbf{m}_i\right)^2}{2} \tag{22}$$

where $w_i = \dfrac{1}{l_1^2 + l_2^2} + \dfrac{1}{l_1'^2 + l_2'^2}$.

Given N point correspondences from the image pair ($\mathbf{m}_i \leftrightarrow \mathbf{m}_i', i = 1 \ldots N$), the search for parameter vector $\mathbf{\Phi}$, which is to be used in $\mathbf{F} = \mathbf{F}(\mathbf{\Phi})$, $\mathbf{H} = \mathbf{H}(\mathbf{\Phi})$, and $\mathbf{H}' = \mathbf{H}'(\mathbf{\Phi})$, becomes a nonlinear optimization problem, that is

$$\min_{\mathbf{F}} \frac{1}{N} \sum_{i=1}^{N} w_i \frac{(\mathbf{m}_i'^T\mathbf{F}\mathbf{m}_i)^2}{2} \tag{23}$$

To simplify the derivation, we can restructure the matrix equation by turning a matrix into a vector. Assume that the 3×3 fundamental matrix in vector form is $\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \mathbf{F}_3 \end{bmatrix}$; then we can use **vec** operator to convert the matrix **F** into a column vector **f** by stacking the columns of **F**, or

$$\mathbf{f} = \mathbf{vec}(\mathbf{F}) = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \\ \mathbf{F}_3 \end{bmatrix} \tag{24}$$

Let symbol $\otimes$ denote Kronecker product; then

$$\mathbf{m}_i'^T\mathbf{F}\mathbf{m}_i = \left(\mathbf{m}_i' \otimes \mathbf{m}_i\right)^T \mathbf{f} \tag{25}$$

Substituting (25) into (23) yields the objective function in vector form. That is

$$\frac{1}{N} \sum_{i=1}^{N} w_i \frac{(\mathbf{m}_i'^T\mathbf{F}\mathbf{m}_i)^2}{2} = \frac{1}{2N} \sum_{i=1}^{N} (w_i^{1/2}\mathbf{m}_i'^T\mathbf{F}\mathbf{m}_i)^2 = \frac{1}{2N} \left\| \mathbf{W}^{1/2}\mathbf{U}_N\mathbf{f} \right\|^2 \tag{26}$$

where **W** is an $N \times N$ diagonal matrix and $\mathbf{U}_N$ is an $N \times 9$ matrix :

$$\mathbf{W} = \mathrm{diag}\begin{bmatrix} w_1 & w_2 & \cdots & w_N \end{bmatrix} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & w_N \end{bmatrix}, \mathbf{U}_N = \begin{bmatrix} (\mathbf{m}'_1 \otimes \mathbf{m}_1)^T \\ (\mathbf{m}'_2 \otimes \mathbf{m}_2)^T \\ \vdots \\ (\mathbf{m}'_N \otimes \mathbf{m}_N)^T \end{bmatrix}.$$

Given these new notation, minimization of the mean-square distance in (23) can now be rewritten as

$$\min_{\mathbf{f}} \ \frac{1}{2N} \left\| \mathbf{W}^{1/2} \mathbf{U}_N \mathbf{f} \right\|^2 \tag{27}$$

The solution for **f** in (27) can be found by any standard nonlinear minimization method. We chose the Levenberg-Marquardt algorithm because of it effectiveness and popularity. Before applying this minimization process, we need to derive the Jacobian matrix of (26), or $\frac{\partial}{\partial \mathbf{\Phi}} (\mathbf{W}^{1/2} \mathbf{U}_N \mathbf{f})$. In order to simplify the computation of this Jacobian matrix, we modify our iteration of the minimization process by using the old **W** value from the previous iteration to compute the new Jocobian matrix. That is,

$$\frac{\partial}{\partial \mathbf{\Phi}_k} \left\{ \mathbf{W}^{1/2}(\mathbf{\Phi}_{k-1}) \mathbf{U}_N \mathbf{f}(\mathbf{\Phi}_k) \right\} = \mathbf{W}^{1/2}(\mathbf{\Phi}_{k-1}) \mathbf{U}_N \frac{\partial}{\partial \mathbf{\Phi}_k} \mathbf{f}(\mathbf{\Phi}_k)$$

Therefore, even though **W** is a function of $\mathbf{\Phi}$, we can treat it as a constant to $\mathbf{\Phi}$ in current iteration. Since factor $\mathbf{W}^{1/2}\mathbf{U}_N$ inside the partial derivative can be treated as constant for $\mathbf{\Phi}$, the Jacobian matrix can then be reduced to a simpler form

$$\frac{\partial}{\partial \mathbf{\Phi}} (\mathbf{W}^{1/2} \mathbf{U}_N \mathbf{f}) = (\mathbf{W}^{1/2} \mathbf{U}_N) \frac{\partial}{\partial \mathbf{\Phi}} \mathbf{f} \tag{28}$$

Therefore, we only need to compute $\frac{\partial}{\partial \mathbf{\Phi}} \mathbf{f}$ for the Jocobian matrix of (26) as follows

$$\frac{\partial}{\partial \mathbf{\Phi}} \mathbf{f} = \frac{\partial}{\partial \mathbf{\Phi}} \begin{bmatrix} -h_4 f \cos\theta + h_7 \sin\theta \\ -h_4 f \sin\theta - h_7 \cos\theta \\ h_4 \\ -h_5 f \cos\theta + h_8 \sin\theta \\ -h_5 f \sin\theta - h_8 \cos\theta \\ h_5 \\ -h_6 f \cos\theta + \sin\theta \\ -h_6 f \sin\theta - \cos\theta \\ h_6 \end{bmatrix}$$

$$= \begin{bmatrix} -h_4 \cos\theta & h_4 f \sin\theta + h_7 \cos\theta & -f\cos\theta & 0 & 0 & \sin\theta & 0 \\ -h_4 \sin\theta & -h_4 f \cos\theta + h_7 \sin\theta & -f\sin\theta & 0 & 0 & -\cos\theta & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -h_5 \cos\theta & h_5 f \sin\theta + h_8 \cos\theta & 0 & -f\cos\theta & 0 & 0 & \sin\theta \\ -h_5 \sin\theta & -h_5 f \cos\theta + h_8 \sin\theta & 0 & -f\sin\theta & 0 & 0 & -\cos\theta \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -h_6 \cos\theta & h_6 f \sin\theta + \cos\theta & 0 & 0 & -f\cos\theta & 0 & 0 \\ -h_6 \sin\theta & -h_6 f \cos\theta + \sin\theta & 0 & 0 & -f\sin\theta & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{29}$$

The Jacobian matrix can be obtained by substituting (29) into (28), which is then used in an iterative process of Levenberg-Marquardt algorithm to find the parameter vector $\mathbf{\Phi}$. This minimization algorithm contains an iterative process, which must start with an initial estimate of the *F matrix*, and the ideal *F matrix* of a pair of rectified images can be used, that is,

$$\hat{\mathbf{F}}_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = \mathbf{F}_\infty$$

A comparison of $\hat{\mathbf{F}}_0$ with the parameterization of $\hat{\mathbf{F}}$ in (17) shows that $\hat{\mathbf{F}}_0$ corresponds to an initial parameter vector of $\mathbf{\Phi}_0 = \begin{bmatrix} f & \theta & h_4 & h_5 & h_6 & h_7 & h_8 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}^T$. An iterative process can now be applied to find the solution of $\mathbf{\Phi}$ after proper stop conditions have been set.

4.4 Homography with minimal geometric distortion

After the parameter vector $\mathbf{\Phi} = \begin{bmatrix} f & \theta & h_4 & h_5 & h_6 & h_7 & h_8 \end{bmatrix}^T$ has been found, the values of the vector can be used to calculate the pair of rectifying homographies shown as below

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix}, \ \mathbf{H}' = \begin{bmatrix} \boldsymbol{cos}\theta & \boldsymbol{sin}\theta & 0 \\ -\boldsymbol{sin}\theta & \boldsymbol{cos}\theta & 0 \\ -f\,\boldsymbol{cos}\theta & -f\,\boldsymbol{sin}\theta & 1 \end{bmatrix} \tag{30}$$

Obviously, the only parameters left to be estimated are $[h_1 \quad h_2 \quad h_3]$. Since this vector does not affect the coordinate of *y*-axis, we can simply set it to $[1 \ 0 \ 0]$ and obtain satisfactory rectifying results. However, to achieve a certain purpose, we can apply specific constraint on the transformed coordinate of *x*-axis and obtain different values for $[h_1 \quad h_2 \quad h_3]$. For example, minimization of equation (11) has been used as an extra constraint to reduce the disparity of *x*-axis between two rectified images. This approach can reduce the range of search for stereo matching and increase the speed on solving the correspondence problem. However, in applications where the *x-axis* disparity should not be modified too much, other constraint can be used for acquiring the values of $[h_1 \quad h_2 \quad h_3]$. One suitable constraint is to keep the aspect ratio of the original image invariant after rectification. We will adopt the idea of shearing transform described in (Loop & Zhang, 1999) to achieve this purpose, and the procedure will be stated below. As will be shown in the experimental section, this approach not only maintains the aspect ratio of the original image but also reduces the overall geometric distortion.

Assume that the parameter vector $\mathbf{\Phi} = \begin{bmatrix} f & \theta & h_4 & h_5 & h_6 & h_7 & h_8 \end{bmatrix}^T$ has been found by following the procedure described in the previous section. Let $[h_1 \quad h_2 \quad h_3] = [1 \quad 0 \quad 0]$; then we have a preliminary solution of homographies

$$\mathbf{H}_0 = \begin{bmatrix} 1 & 0 & 0 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix}, \ \mathbf{H}'_0 = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ -f\cos\theta & -f\sin\theta & 1 \end{bmatrix}$$

To keep the aspect ratio invariant after rectification, these two homographies can further be combined with the shearing transform defined below

$$\mathbf{H}_s = \begin{bmatrix} a & b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ \mathbf{H}'_s = \begin{bmatrix} a' & b' & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{31}$$

The final homographies for the rectification can then be written as

$$\mathbf{H} = \mathbf{H}_s\mathbf{H}_0 = \begin{bmatrix} a & b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\mathbf{H}_0 \ , \ \mathbf{H}' = \mathbf{H}'_s\mathbf{H}'_0 = \begin{bmatrix} a' & b' & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\mathbf{H}'_0 \tag{32}$$

Using $\mathbf{H}_0$ and $\mathbf{H}'_0$ alone can rectify the image pair such that the coplanar condition is satisfied; however, combined with $\mathbf{H}_s$ and $\mathbf{H}'_s$ respectively, we can further improve the appearance of the final rectification results. A detailed description of the shearing transform can be found in (Loop & Zhang, 1999), but its adaptation to our usage will be briefly described below.

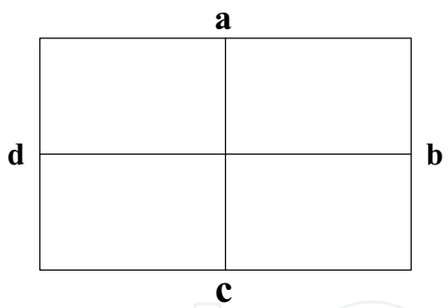


Figure 3. Center points used for computing shearing transform matrices

For a given image with width w and height h, coordinates of the midpoints on its four boundaries are shown in Fig. 3 and can be represented as

$$\mathbf{a} = \begin{bmatrix} \frac{w-1}{2} & 0 & 1 \end{bmatrix}^T, \mathbf{b} = \begin{bmatrix} w-1 & \frac{h-1}{2} & 1 \end{bmatrix}^T, \mathbf{c} = \begin{bmatrix} \frac{w-1}{2} & h-1 & 1 \end{bmatrix}^T, \mathbf{d} = \begin{bmatrix} 0 & \frac{h-1}{2} & 1 \end{bmatrix}^T$$

The two central lines are expressed as

$$\mathbf{x} = \mathbf{b} - \mathbf{d} \ , \ \mathbf{y} = \mathbf{c} - \mathbf{a}$$

Let $\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{c}}, \hat{\mathbf{d}}$ be the coordinates of these four midpoints after transformation by $\mathbf{H}_0$, or

$$\hat{\mathbf{a}} = \begin{bmatrix} \hat{a}_u \\ \hat{a}_v \\ \hat{a}_w \end{bmatrix} = \mathbf{H}_0\mathbf{a} \xrightarrow{normalize} \begin{bmatrix} \hat{a}'_u \\ \hat{a}'_v \\ 1 \end{bmatrix},\ \hat{a}'_u = \frac{\hat{a}_u}{\hat{a}_w}\ ,\ \hat{a}'_v = \frac{\hat{a}_v}{\hat{a}_w} \tag{33}$$

The other three points ($\hat{\mathbf{b}}$, $\hat{\mathbf{c}}$, $\hat{\mathbf{d}}$) can be found by the same way, and the two central lines after rectification become

$\hat{\mathbf{x}} = \hat{\mathbf{b}} - \hat{\mathbf{d}} = \begin{bmatrix} x_u & x_v & 0 \end{bmatrix}^T$, $\hat{\mathbf{y}} = \hat{\mathbf{c}} - \hat{\mathbf{a}} = \begin{bmatrix} y_u & y_v & 0 \end{bmatrix}^T$

The perspective component of $\mathbf{H}_0$ causes the projective rectification to generate distortion, and the shearing transform $\mathbf{H}_s$, which is used to reduce the distortion, can be found by satisfying the following two constraints:

1. Orthogonal:: $$\mathbf{x}^T\mathbf{y} = (\mathbf{H}_s\hat{\mathbf{x}})^T(\mathbf{H}_s\hat{\mathbf{y}}) = 0 \tag{34}$$

2. Invariant aspect ratio: $$\frac{\mathbf{x}^T\mathbf{x}}{\mathbf{y}^T\mathbf{y}} = \frac{(\mathbf{H}_s\hat{\mathbf{x}})^T(\mathbf{H}_s\hat{\mathbf{x}})}{(\mathbf{H}_s\hat{\mathbf{y}})^T(\mathbf{H}_s\hat{\mathbf{y}})} = \frac{\mathrm{w}^2}{\mathrm{h}^2} \tag{35}$$

where $\mathbf{H}_s = \begin{bmatrix} a & b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Expanding equations (34), (35) and solving the quadratic polynomials of a and b based on the techniques similar to that were described in (Loop & Zhang, 1999), we obtain

$$a = \frac{\mathrm{h}^2x_v^2 + \mathrm{w}^2y_v^2}{\mathrm{hw}(x_vy_u - x_uy_v)}\ ,\ b = \frac{\mathrm{h}^2x_ux_v + \mathrm{w}^2y_uy_v}{\mathrm{hw}(x_uy_v - x_vy_u)} \tag{36}$$

Substituting a, b into equation (31) yields shearing transform matrix $\mathbf{H}_s$. In order to make all the rectified pixels appear within the visible range, a must be positive. If a is a negative value, then both a and b are multiplied by -1. Elements a' and b' of matrix $\mathbf{H}'_s$ can be found by the same way. After $\mathbf{H}_s$ and $\mathbf{H}'_s$ were obtained, the final homographies used for rectification with minimal distortion become $\mathbf{H} = \mathbf{H}_s\mathbf{H}_0$, $\mathbf{H}' = \mathbf{H}'_s\mathbf{H}'_0$. They then can be used for resampling to complete the process of projective rectification.

5. Results and discussion

To evaluate performance of the proposed method, several indoor image pairs acquired from the INRIA web site and CMU (please see Acknowledge section) were tested. For each image pair, a set of ten point correspondences were selected from each image and used to compute the *H matrices* for rectification. Selection of these point correspondences takes a semi-automatic approach to avoid inaccuracy. That is, each point is chosen approximately by hand, and then a precise (subpixel accuracy) feature point close-by is then detected by Harris corner finder (Harris, 1998) inside a local search window. This approach not only makes the selection very easy, it also greatly improves the accuracy. To achieve the best results, these points are chosen to be evenly distributed over the entire image area.

5.1 Visual evaluation of rectifying results

The original and rectified image pairs are shown in Figures 4~8 for visual comparison, where the top is the original image pair, the middle and the bottom are the rectified results of the method proposed and the Hartley's method, respectively. To achieve a more robust estimation of **F** matrix in Hartley's method, RANSAC (Torr & Murry, 1997) is used. RANSAC calculates for each **F** the number of inliers, in which the chosen **F** is the one that maximizes it. Once the outliers are removed, **F** is recalculated with the aim of obtaining a better estimation. The solution for the pair of homogphies in the proposed method can be found in less than 100 iterations by Levenberg-Marquardt algorithm, and the re-sampling can be done in real time by look up table. As shown in these figures, all the re-sampled image pairs are properly rectified by the proposed method with minimal geometric distortions. In order to make visual evaluation of the rectified results more convenient, four horizontal lines are added to identify the difference of y-disparity before and after rectification.

In all the figures, it's obvious that our proposed method has less geometric distortion than that of Hartley's method. In Fig. 4, the Hartley's method greatly distorts the right image in the process of minimizing x disparity. Our method shows its capability in maintaining the angle and aspect ratio of the objects in the scene. However, the Hartley's method is not able to keep these properties invariant. The main reason for the distortion after rectification is because the image content contains a variety of depth values and therefore many different amounts of y-disparity. When the rectification algorithm tries to minimize the disparity all over the image region, distortion occurs. Overall, the proposed method keeps the objects in better shape than that of the Hartley's method. The greater geometric distortion of the Hartley's method is due to its minimization of the x-disparity.

5.2 Quantitative evaluation of rectifying results

In addition to the above visual comparisons, quantitative evaluation based on the changes in y disparity is also conducted as follows. The (x, y) coordinates for the ten chosen point correspondences in one image pair (Balmouss) before and after rectification by the proposed method are shown in Tables 1 for numerical evaluation. Obviously, the x disparities for the selected points have not changed too much after rectification. To estimate the accuracy of the rectification process numerically, we define the mean of the absolute difference (MAD) of y coordinate, $\overline{|\Delta y|}$, for the original and rectified image pair as

$$\overline{|\Delta y|} = \begin{cases} \varepsilon_{org} = \frac{1}{N}\sum_{i=1}^{N} |\,(\mathbf{m}_i)_y - (\mathbf{m}'_i)_y\,|: & \text{original} \\ \varepsilon_{rec} = \frac{1}{N}\sum_{i=1}^{N} |\,(\mathbf{H}\mathbf{m}_i)_y - (\mathbf{H}'\mathbf{m}'_i)_y\,|: & \text{rectified} \end{cases} \tag{37}$$

where $(\cdot)_y$ indicates the y coordinates and $(\mathbf{m}_i, \mathbf{m}'_i)$ represents the coordinates of the i^{th} pair of the point correspondences.

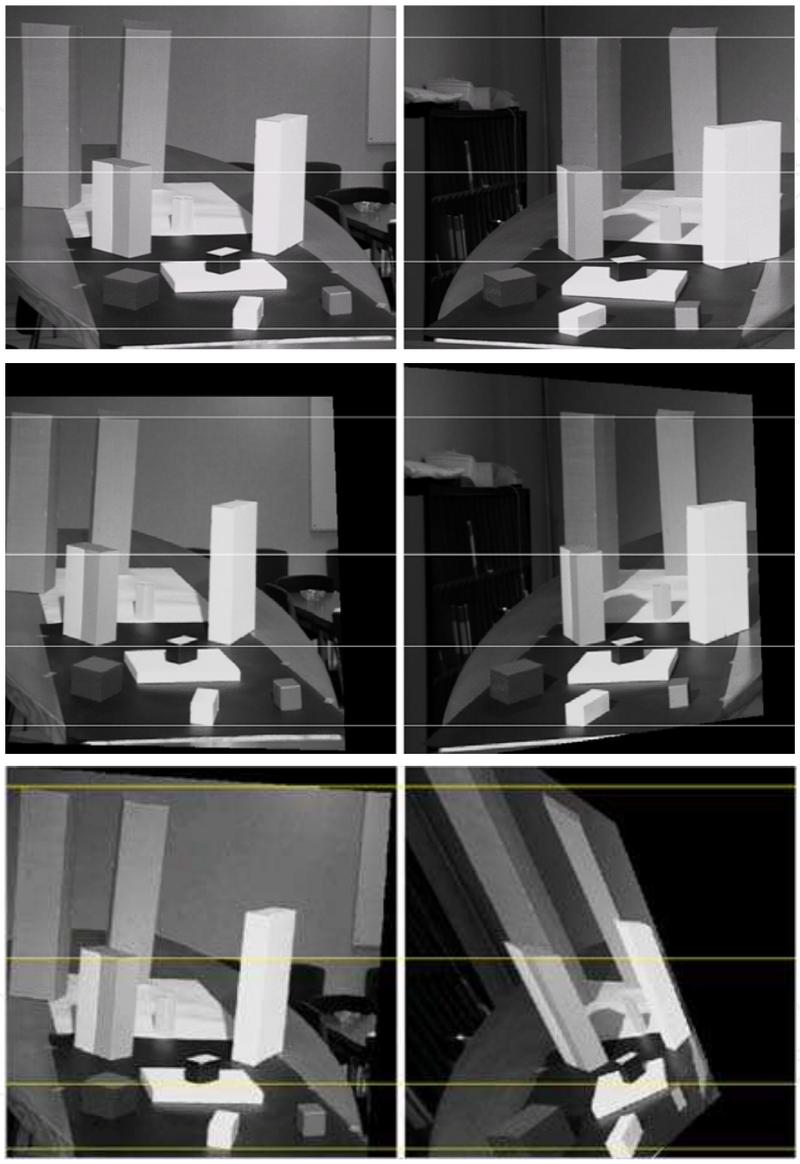


Figure 4. **Aout** image pair of INRIA. (a)Top row: original images. (b)Middle row: rectified images using the proposed method. (c)Bottom row: rectified images using Hartley's method. The proposed method has much lower visual distortion.

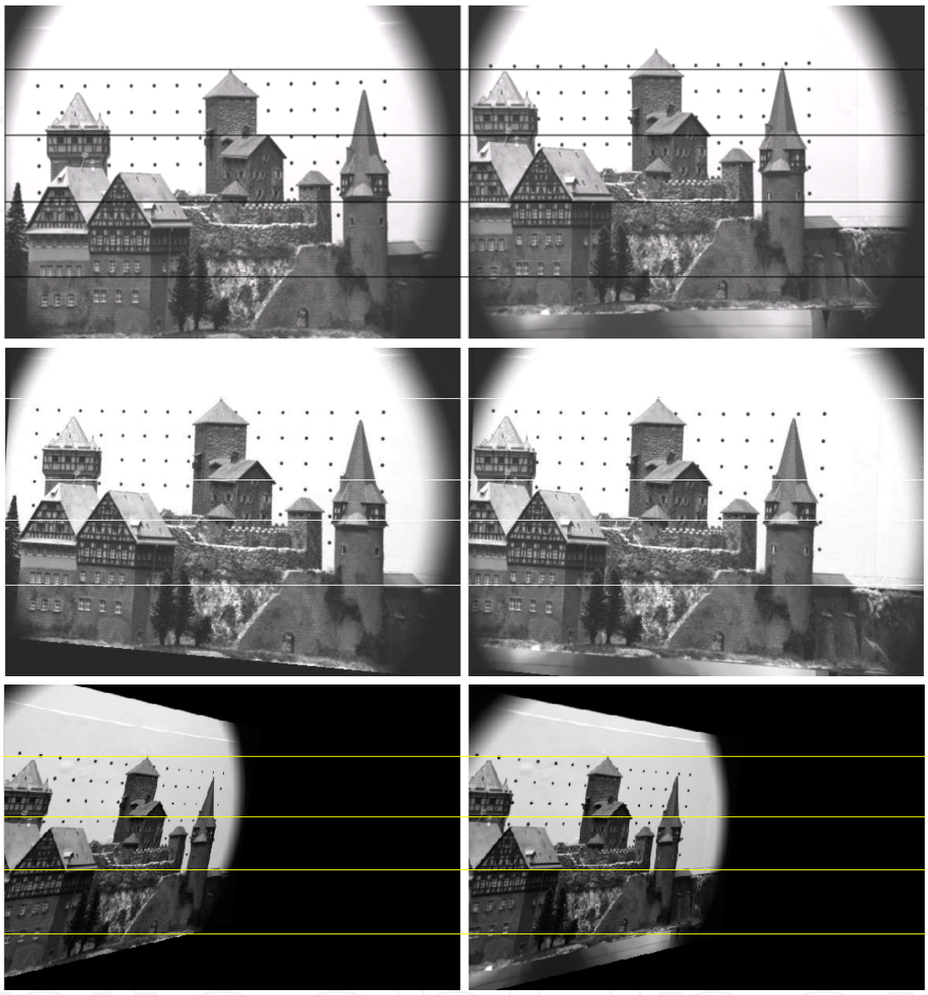


Figure 5. Rectifying results of the **castle** image from CMU/CIL (vary large y-disparity)

(a) Top row: original images.

(b) Middle row: rectified images using the proposed method.

(c) Bottom row: rectified images using Hartley's method.

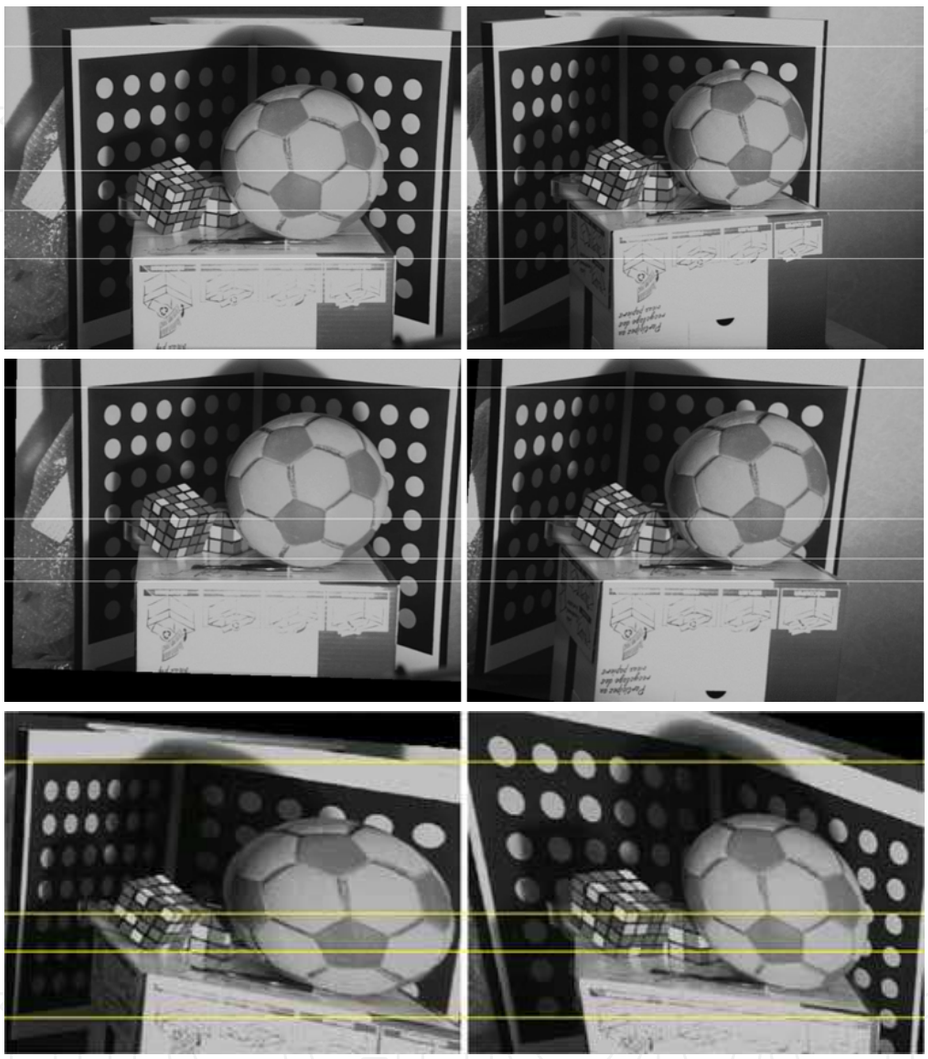


Figure 6. Rectifying results of the **BalMouss** image pair

(a) Top row: original pair of images.

(b) Middle row: rectified images using the proposed method.

(c) Bottom row: rectified images using Hartley's method

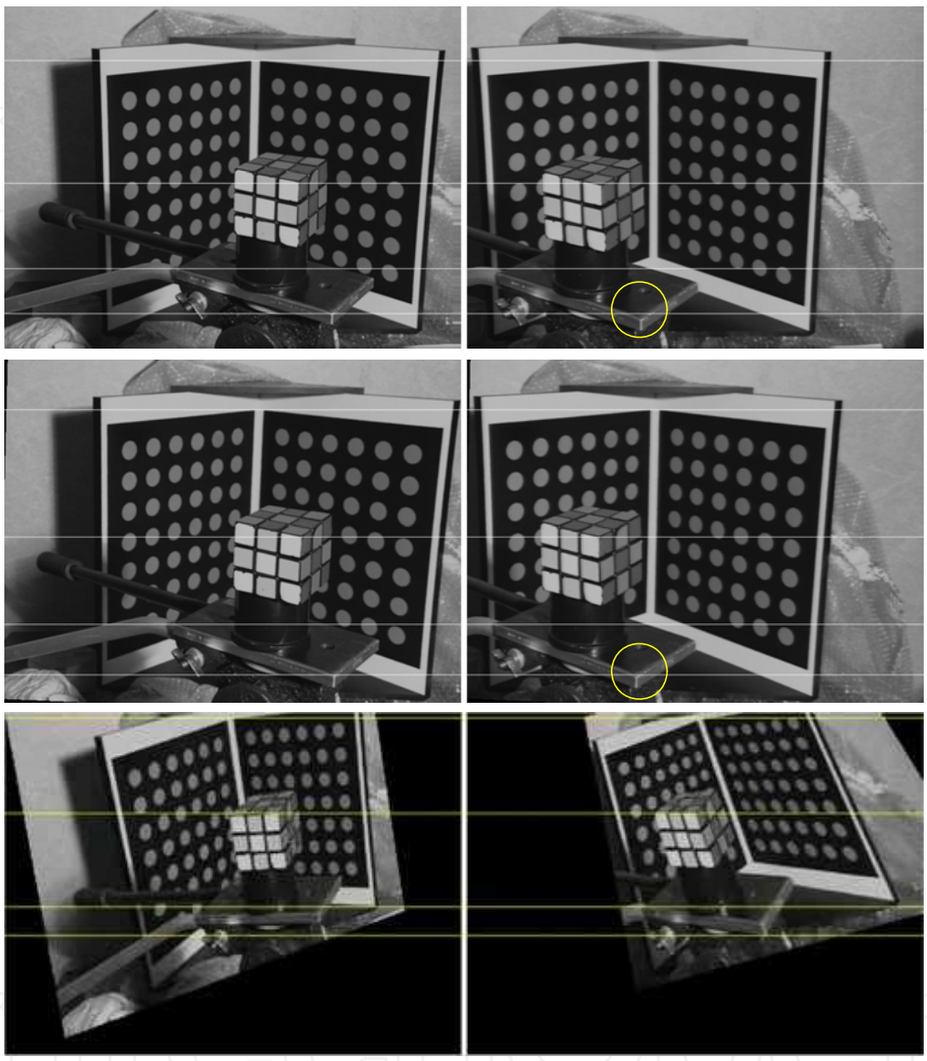


Figure 7. Rectifying results of **Rubik** image pair from INRIA

(a) Top row: original images.

(b) Middle row: rectified images using the proposed method.

(c) Bottom row: rectified images using Hartley's method

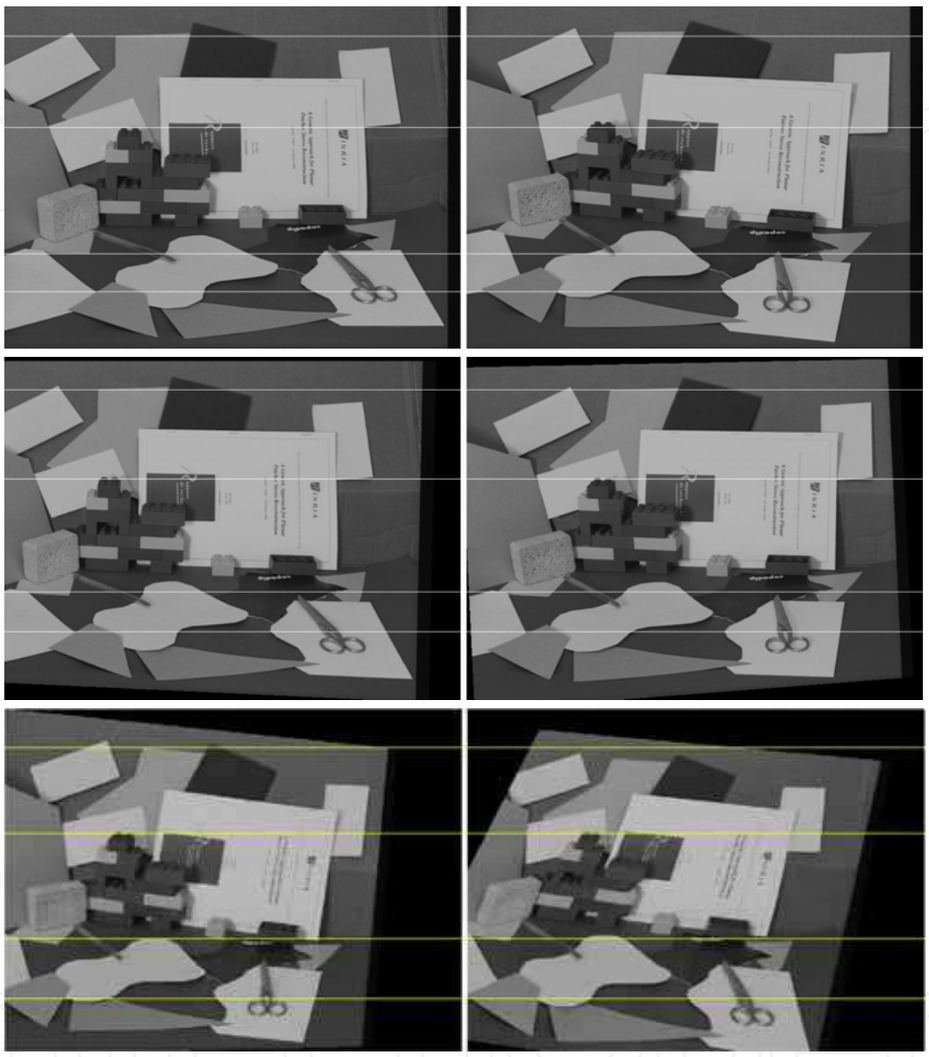


Figure 8. Rectifying results of **Tot** image pair from INRIA

(a) Top row: original images.

(b) Middle row: rectified images using the proposed method.

(c) Bottom row: rectified images using Hartley's method

A desirable property of the method is to *minimize* the $\overline{|\Delta y|}$, but keep the value of $\overline{|\Delta x|}$ *invariant* after rectification. When these ten chosen point correspondences were used for $\overline{|\Delta y|}$ computation, the results before and after rectification for the above six image pairs are listed in Table 2. If an image pair is ideally rectified, $\overline{|\Delta y|}$ of the point correspondences after rectification should be zero. As can be seen from the table, for all the image pairs, values of $\overline{|\Delta y|}$ after rectification of our proposed method (ε_{rec}) are all less than 1 pixel. They are greatly reduced compared with the values before rectification (ε_{org}), indicating the effectiveness of the proposed method. However, some image pairs can not be satisfactorily rectified by the Hartley's method.

Before rectification

Left image	x-axis	127	136	231	253	411	271	275	361	528	736
	y-axis	91	533	336	312	65	321	481	349	499	93
Right image	x-axis	55	64	208	235	255	255	297	320	508	587
	y-axis	77	501	302	277	49	286	428	307	426	69

After rectification

Left image	x-axis	127.37	136.81	233.38	256.66	419.84	274.49	278.59	368.33	543.28	766.42
	y-axis	77.376	509.38	318.49	295.32	53.493	304.77	462.15	333.97	487.31	82.733
Right image	x-axis	54.952	63.16	217.2	248.1	271.86	270.85	317.76	345.53	576.92	682.91
	y-axis	77.7	509.42	318.34	294.79	53.058	305.45	462.12	333.87	487.32	82.917

Table 1. Coordinates of point correspondences before and after rectification for **Balmouss** image pair based on the proposed method.

Image name / MAD_*y*	Aout Image pair	Castle image pair	Balmouss Image pair	Rubik image pair	Tot image pair
ε_{org} (pixel)	13.6	26.7084	35.8598	12.8	13.9577
ε_{rec} (pixel)	0.6322	0.5084	0.2477	0.6389	0.2779
ε_{rec_H} (pixel)	14.9539	0.4731	21.8769	14.2357	4.8668

Table 2. MAD of *y*-coordinate before (ε_{org}) and after rectification with the proposed (ε_{rec}) or Hartley's method (ε_{rec_H}). Evaluation based on 10 selected point correspondences.

6. Comparisons and discussion

Compared with the results presented in Hartley (Hartley, 1999), our new method has the following advantages:

1. Our method has much less geometric distortion visually, even when the image pair has large difference of viewpoints (Fig. 4). Further comparisons with the Hartley's method on geometric distortion are presented in Figs. 5-8 using four other image pairs.
2. The new method avoids using the similar constraint shown in (11) in order to obtain a unique solution. This extra x-axis disparity minimization step, which is used in deriving the Hartley's method, will be unreasonable if the rectified result is used for stereoscopic viewing. Instead of minimizing x-axis disparity, shearing transform is used in our algorithm to preserve aspect ratio and reduce geometric distortion.
3. Solving the rectification problem directly without first computing the **F** matrix makes the proposed method avoid the problem of selecting proper method for **F** matrix estimation.
4. Initial value of the nonlinear solution by iteration is much easier to set. The initial parameter vector $\mathbf{\Phi}_0$ is simply set to $\begin{bmatrix}1 & 0 & 0 & 1 & 0 & 0 & 0\end{bmatrix}^T$.

We use the same initial guess of the optimization process for all the images tested and the solutions always converge. The iteration stops after the error is smaller than a preset threshold. Although we are not sure if the true minimum is obtained, the rectified results show its robustness, even for image pair with very different view like Fig. 4. Since we are not looking for the true minimum to obtain an optimal solution, the convergence towards the true minimum is not guaranteed. If further improvement is needed, some approaches which can avoid local minimum might be taken.

Most of the projective rectification methods proposed all base their algorithms on an estimation of the *F* matrix. However, as has been stated in (Zhang, 1998), the *F* matrix estimator has its own uncertainty. Our approach, similar to the IT method (Isgrò & Trucco, 1999), avoids *F* matrix estimation procedure and obtains homographies directly. Furthermore, it improves on the Hartley's method and obtains rectifying results with reduced geometric distortion.

7. Conclusion

This chapter presented a new way of parameterizing the homography, which leads to a new approach of projective rectification for stereo images. Compared with the previous works, the novelty of this new algorithm is that it uses mean-square distance as minimization criterion which has more well-defined geometric meaning. Furthermore, instead of putting constraint on x-axis disparity, we use shearing transform to achieve a single solution for the projective rectification problem, and greatly reduce the geometric distortion. Visual inspection and quantitative evaluation of the rectification results show the accuracies of the proposed method and its low geometric distortion. Experiments on different types of image pairs with various *y*-disparity values have been conducted, and the results show that the proposed method can effectively reduce the geometric distortion.

8. Acknowledgement

The authors gratefully acknowledge the sources of some of the stereo image pairs used in this paper, which are kindly available from INRIA and CMU/CIL under Copyright.
INRIA: http://www-rocq.inria.fr/mirages/SYNTIM_OLD/analyse/paires-eng.html
CMU/CIL: http://www-2.cs.cmu.edu/afs/cs.cmu.edu/project/cil/www/cil.html

The Matlab codes of rectification implementation offered by (Hartley, 2004) for comparison is greatly appreciated.

9. References

Al-Shalfan, KA.; Haigh, JGB. & Ipson, SS. (2000). Direct algorithm for rectifying pairs of uncalibrated images, *Electronics Letters*, **36**(5), pp. 419-420.

Ayache, N. & Hansen, C. (1988). Rectification of images for binocular and trinocular stereovision, *IEEE ICPR*, pp. 11-16, Nov , Rome, Italy.

Ayache, N. & Lustman, F.(1991). Trinocular stereo vision for robotics. *IEEE Trans. on PAMI*, 13, Jan, pp. 73-85, ISSN: 0162-8828.

Faugeras, O. (1993). *Three-Dimensional Computer Vision*, MIT Press, Cambridge, MA, ISBN: 0387151192.

Fusiello, A.; Trucco, E. & Verri, A. (2000). A compact algorithm for rectification of stereo pairs**.** *Machine Vision and Applications*, **12**, July, pp. 16–22, ISSN: 0932-8092.

Gluckman, J. & Nayar, S. (2001). Rectifying Transformations That Minimize Resampling Effects, *IEEE CVPR'01*, pp. 111-117, Dec, Kauai, HI, USA.

Harris, C. & Stephens, M. (1998). A Combined Corner and Edge Detector, *Alvey Vision Conf,* pp. 147-151.

Hartley, R. (1997). In Defense of the Eight-Point Algorithm, *IEEE Transactions on PAMI*, **19**, pp. 580-593

Hartley, R. (1999). Theory and Practice of Projective Rectification. *International Journal of Computer Vision*, **35** , pp. 115-127.

Hartley, R (2004). http://www.robots.ox.ac.uk/~vgg/hzbook/code

Isgrò, F. & Trucco, E. (1999). On Robust Rectification for Uncalibrated Images, *IEEE International Conference on Image Analysis and Processing*, pp. 297-302, Sep.

Loop, C. & Zhang, Z. (1999). Computing Rectifying Homographies for Stereo Vision, *IEEE CVPR'99*, pp. 125-131, June, Fort Collins, CO, USA.

Papadimitriou, DV. & Dennis, TJ. (1996). Epipolar line estimation and rectification for stereo image pairs, *IEEE Transactions on Image Processing*, **5**(4), pp. 672-676.

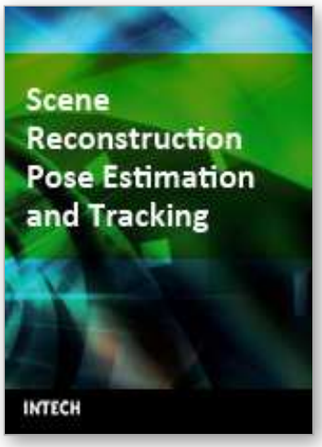
Pollefeys, M.; Kock, R. & Gool, LV. (1999). A Simple and Efficient Rectification Method for General Motion, *Proc. ICCV*, pp. 496-501, Corfu, Greece.

Robert, L.; Zeller, C. & Faugeras, O. et al., (1997). Applications of non-metric vision to some visually-guided robotics tasks, In : *Visual Navigation: From Biological Systems to Unmanned Ground Vehicles (Vol. II of Advances in Computer Vision, Lawrence Erlbaum Associates)*, Aloimonos Y, ed.

Slama, C. (1980). *Manual of Photogrammetry*, American Society of Photogrammetry, ISBN: 0937294012.

Torr, P.H.S. & Murry D.W. (1997). The development and comparison of robust methods for estimating the fundamental matrix, *International Journal of Computer Vision* 24 (3) , pp.271–300.

Zhang, Z. (1998). Determining the Epipolar Geometry and its Uncertainty : A Review, *International Journal of Computer Vision*, **27**, pp. 161-195.



Scene Reconstruction Pose Estimation and Tracking

Edited by Rustam Stolkin

ISBN 978-3-902613-06-6

Hard cover, 530 pages

Publisher I-Tech Education and Publishing

Published online 01, June, 2007

Published in print edition June, 2007

This book reports recent advances in the use of pattern recognition techniques for computer and robot vision. The sciences of pattern recognition and computational vision have been inextricably intertwined since their early days, some four decades ago with the emergence of fast digital computing. All computer vision techniques could be regarded as a form of pattern recognition, in the broadest sense of the term. Conversely, if one looks through the contents of a typical international pattern recognition conference proceedings, it appears that the large majority (perhaps 70-80%) of all pattern recognition papers are concerned with the analysis of images. In particular, these sciences overlap in areas of low level vision such as segmentation, edge detection and other kinds of feature extraction and region identification, which are the focus of this book.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen