Rectifying Transformations That Minimize Resampling Effects *

Joshua Gluckman

Dept. of Computer Science
Polytechnic University
Brooklyn, NY 11201
jgluckma@poly.edu

Shree K. Nayar

Dept. of Computer Science
Columbia Univerisity
New York, NY 10027
nayar@cs.columbia.edu

Abstract

Image rectification is the process of warping a pair of stereo images in order to align the epipolar lines with the scan-lines of the images. Once a pair of images is rectified, stereo matching can be implemented in an efficient manner. Given the epipolar geometry, it is straightforward to define a rectifying transformation, however, many transformations will lead to unwanted image distortions. In this paper, we present a novel method for stereo rectification that determines the transformation that minimizes the effects of resampling that can impede stereo matching. The effects we seek to minimize are the loss of pixels due to under-sampling and the creation of new pixels due to over-sampling. To minimize these effects we parameterize the family of rectification transformations and solve for the one that minimizes the change in local area integrated over the area of the images.

1 Introduction

Stereo matching is the process of finding corresponding points in a pair of images. Given a point in one image its matching point must lie on an epipolar line in the other image. This is the well known *epipolar constraint*. The location of the epipolar lines is determined by the epipolar geometry, thus knowledge of the epipolar geometry reduces stereo matching from a 2-D to a 1-D search. When the images are obtained from an identical pair of cameras pointing in the same direction, known as a rectilinear stereo rig, the epipolar lines coincide with the scan-lines of the images. For stereo views in this configuration, stereo matching can be implemented efficiently for several reasons:

- Because the epipolar line for a point is given by the scan-line of that point, the computation of the locations of the epipolar lines can be avoided.
- Since there is no relative rotation between the images, the matching windows do not need to be rotated for robust matching.
- During matching, computational redundancies within and between epipolar lines greatly reduce the number of operations needed to compute correspondences between image points.

*This work in part was supported by a National Science Foundation ITR Award No. IIS-00-85864.

When the epipolar geometry is not in this form, the stereo images need to be warped to make corresponding epipolar lines coincident with the scan-lines, a process termed *image rectification*. Most stereo algorithms assume images are in this configuration, thus rectification is a pre-requisite for stereo matching.

Rectification is achieved by applying a perspective transformation to each image that projects the images onto a plane parallel to the line connecting the centers of projection of the two views. Because there are an infinite number of rectifying planes of which many will lead to unwanted image distortions, careful consideration must be given to the rectification process.

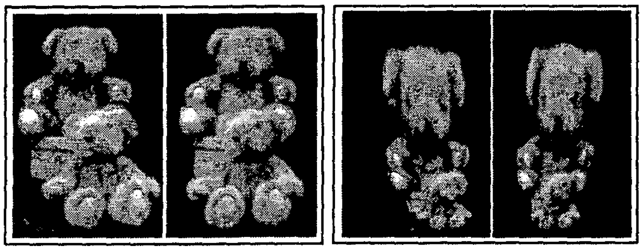
1.1 Previous Work

Traditional methods for rectification require three dimensional information such as the relative orientation of the stereo rig or the camera projection matrices. These methods choose a plane "close" to the image planes of the two cameras without regard to the effects of rectification on the images. Examples include [2] and [3] that use the plane containing the intersection of the two image planes, and [6] that considers the special case of a verged stereo rig. When the cameras are close to rectilinear these methods may suffice, however, for more general camera positions there are more "optimal" methods.

More recently, techniques have been proposed that rectify directly from knowledge of the epipolar geometry without the need for camera calibration or projection matrices. What distinguishes these methods from each other is the metric that is used to find the "best" rectifying transformation. Robert *et al.* [8], the first to consider the effects of rectification on images, attempt to reduce the amount of distortion by finding the rectification transform that is closest to preserving orthogonality about the image centers. Hartley [4] suggests using the transformation that minimizes the range of disparity between the two images in order to minimize differences between the images. Most similar to our work is that of Loop and Zhang [5] who consider the effects of rectification throughout the image and find the rectifying perspective warp that is "closest to affine" over the area of the images. However, their approach does not consider the effects of scale, aspect-ratio, and skew because these are invariant to affine transformations. Furthermore, it is not clear that the measure "closest to affine" is a good one.

In structure from motion, when the direction of heading

0-7695-1272-0/01 $10.00 © 2001 IEEE



(a) (b)

Figure 1: Not all rectified images are the same. (a) On the left is a pair of rectified stereo images. (b) The images on the right are also rectified, however, due to a poor choice of the rectifying transformation a large amount of under-sampling and over-sampling exists which makes stereo matching difficult.

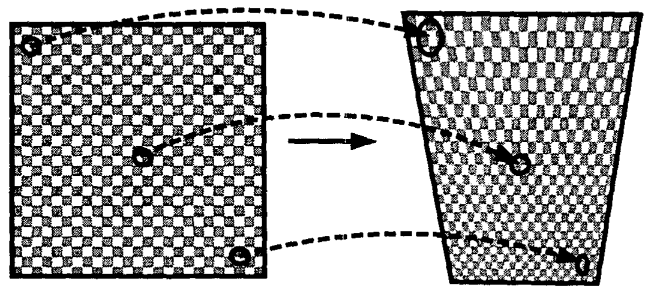


Figure 2: When a perspective transformation is applied to an image the effect on the sampling is modeled by the change in local area. Depending on the location in the image the local area may shrink causing a loss of pixels, grow causing the creation of pixels, or remain constant.

can be located within the image, it is necessary to apply a non-perspective warp in order to rectify the entire image. Rectification under these circumstances has been discussed in [9] and [7].

1.2 Our Approach

In this paper, we pose the stereo rectification problem as finding the transformation that best preserves the sampling of the original stereo pair. Ideally, we would like each pixel in the unrectified image to map to a single pixel in the rectified image. However, when a perspective warp is applied to an image some portions of the image can increase in scale causing the creation of new pixels while other portions can decrease in scale causing the loss of pixels. Both of these effects, which we call over-sampling and under-sampling respectively, can impede stereo matching. Over-sampling can smooth out image texture that is needed for robust stereo matching while under-sampling causes both aliasing and a loss of information.

In our approach, we first model the loss and creation of pixels by the change in local area. Next, we derive a parameterization of the family of perspective transformations that rectify a given stereo pair. Then, we solve for the transforms that minimize the change in local area integrated over the area of the images.

2 Preserving Local Sampling

Figure 1 demonstrates the effect of a rectifying warp that is chosen poorly, causing both the creation and loss of pixels. In order to find the rectification transformation that best preserves the sampling of the stereo images we need a metric for measuring the amount of over-sampling and under-sampling caused by an image transformation.

Both the creation and loss of pixels at a point in the image is modeled by the change in local area of a small patch around the point before and after the transformation is applied (see figure 2). The change in local area is given by the determinant of the Jacobian of the transformation [1]. When the determinant of the Jacobian is one, the area remains constant. As the determinant approaches zero the local area vanishes and thus the pixel at that point is effectively lost. For values greater than unity the local area increases and the number of pixels created is proportional to the change. Note that when the transformation is a translation, a 2-D rotation, a skew or a change of aspect ratio, the local area does not change. For the case of translation and rotation there is a one-to-one mapping between pixels before and after the transformation, therefore no pixels are lost and no new pixels are created. However, due to the finite size of pixels a change of aspect ratio will effect the sampling by introducing new pixels in the direction of increasing scale and destroying pixels in the direction of decreasing scale. Likewise, skew will introduce resampling effects by causing aliasing. We will address this problem later.

For the purposes of rectification, we are interested in perspective transformations. When a perspective transformation is applied to an image, a point (x, y) is mapped to the point $(\hat{x}, \hat{y})$ as

$$\hat{x} = \frac{p_1 x + p_2 y + p_3}{p_7 x + p_8 y + p_9} \text{ and } \hat{y} = \frac{p_4 x + p_5 y + p_6}{p_7 x + p_8 y + p_9}, \tag{1}$$

where the p_i are the parameters of the transformation. The Jacobian is obtained by taking partial derivatives of the above equations with respect to x and y:

$$\frac{\partial(\hat{x}, \hat{y})}{\partial(x, y)} = \begin{pmatrix} \frac{\partial \hat{x}}{\partial x} & \frac{\partial \hat{x}}{\partial y} \\ \frac{\partial \hat{y}}{\partial x} & \frac{\partial \hat{y}}{\partial y} \end{pmatrix}. \tag{2}$$

Then, the determinant of the Jacobian of a perspective transformation is

$$\frac{p_9(p_5 p_1 - p_4 p_2) + p_8(p_4 p_3 - p_1 p_6) + p_7(p_2 p_6 - p_5 p_3)}{(p_7 x + p_8 y + p_9)^3}. \tag{3}$$

Since the change in local area of a perspective transformation is dependent on the location (x, y) in the image, the determinant can be less than one in some places and greater

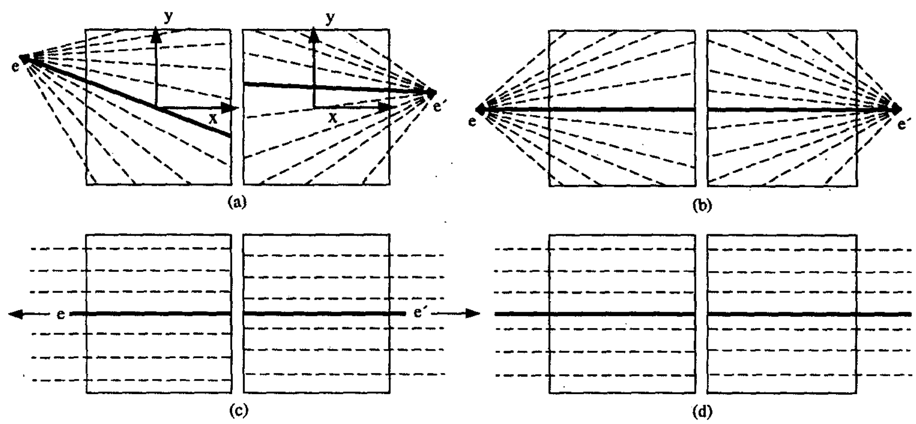


Figure 3: In (a) we have the original unrectified images with epipoles $\mathbf{e}$ and $\mathbf{e}'$ and one pair of corresponding epipolar lines marked in bold. Rectification is composed of three steps. In the first step, shown in (b), the images are rotated and translated so that the marked lines are horizontal and matched on the x-axis. (c) The epipolar lines are made horizontal by projecting the epipoles to infinity. (d) Finally, a transformation is applied which matches the epipolar lines on the same scan-lines.

than one in others. In order to find the rectification transformation that best preserves the local area over the entire image, an error metric is needed. We use the square of the difference between the determinant and one and integrate over the width w and the height h of the image to obtain

$$\int_{-\frac{h}{2}}^{\frac{h}{2}} \int_{-\frac{w}{2}}^{\frac{w}{2}} \left(\det \frac{\partial(\hat{x}, \hat{y})}{\partial(x, y)} - 1 \right)^2 dx\, dy. \tag{4}$$

This metric penalizes the loss of a pixel the same as the creation of a pixel as long as the determinant is less than two. Using the absolute value is an alternative and avoids biasing the error when the determinant is greater than two, but at the cost of integrability. Note that when the scene of interest lies in a particular region of the image the integration can be restricted to that region.

3 Rectifying Perspective Transformations

In this section we derive the entire class of perspective transformations that rectify a stereo image. Then, we will show how to find the transform within that class that minimizes the measure (4) for both stereo images. We assume the epipolar geometry of the stereo images $\mathcal{I}$ and $\mathcal{I}'$ is known and given by the fundamental matrix $\mathbf{F}$. There are many methods for estimating $\mathbf{F}$ from a pair of images (see for example [10]). If $\mathbf{x}$ and $\mathbf{x}'$ in $\mathcal{I}$ and $\mathcal{I}'$ are corresponding image projections of the same scene point then $\mathbf{Fx}$ and $\mathbf{F}^T\mathbf{x}'$ are *corresponding* epipolar lines. The epipolar lines for each image intersect at the epipoles $\mathbf{e}$ and $\mathbf{e}'$ which are the solutions to $\mathbf{Fe} = 0$ and $\mathbf{F}^T\mathbf{e}' = 0$.

In general, the epipoles lie in the image plane and therefore epipolar lines do not run along the scan-lines. To rectify, perspective transformations $\mathbf{P}$ and $\mathbf{P}'$ are applied to the images respectively. These transformations must both project the epipoles to infinity in the direction of the scan-lines and ensure that corresponding lines are on the same scan-line (see figure 3). First, we will derive the constraints on $\mathbf{P}$ and $\mathbf{P}'$ that bring the epipolar lines into a horizontal configuration. Then, we will give the constraints between $\mathbf{P}$ and $\mathbf{P}'$ that ensure corresponding epipolar lines are on the same scan-line.

We place the origin of the coordinate system at the center of each image with the x-axis running along the scan-lines. To simplify, we put the fundamental matrix into a canonical form by applying a rotation and translation so that one pair of corresponding epipolar lines coincides with the x-axis (see figure 3(b)). Note that applying a pure 2-D rotation or translation to the images does not affect the sampling because they are one-to-one mappings. We match the pair of epipolar lines by applying rotation $\mathbf{R}$ to image $\mathcal{I}$ so that the epipolar line through the center of the image is aligned with the x-axis. Next a rotation $\mathbf{R}'$ is applied to image $\mathcal{I}'$ to make the corresponding epipolar line parallel. Then a translation $\mathbf{T}'$ is used to make this epipolar line coincident with the x-axis. Once rotated and translated, we obtain a new fundamental matrix:

$$\hat{\mathbf{F}} = (\mathbf{T}'\mathbf{R}')^{-T}\, \mathbf{F}\mathbf{R}^T. \tag{5}$$

Because the epipolar line of the origin is the x-axis for both the left and right image, it is straightforward to show that the new fundamental matrix must be of the form

$$\hat{\mathbf{F}} = \begin{pmatrix} 0 & f_2 & 0 \\ f_4 & f_5 & f_6 \\ 0 & f_8 & 0 \end{pmatrix}. \tag{6}$$

Now the epipoles are on the x-axis and are expressed in projective coordinates as $\mathbf{e} = \left[1, 0, -\frac{f_4}{f_6}\right]^T$ and $\mathbf{e}' = \left[1, 0, -\frac{f_2}{f_8}\right]^T$. Next, we make all the epipolar lines parallel (see figure 3(c)) by imposing the following constraints on $\mathbf{P}$ and $\mathbf{P}'$:

- Fix the origins of $\mathcal{I}$ and $\mathcal{I}'$ such that $(0,0)$ maps to $(0,0)$.
- Project the epipoles to infinity such that $\mathbf{e}$ and $\mathbf{e}'$ map to $\mathbf{i} = [1,0,0]^T$.

These conditions are satisfied if $\mathbf{P}$ and $\mathbf{P}'$ are of the form

$$\mathbf{P} = \begin{pmatrix} p_1 & p_2 & 0 \\ 0 & p_5 & 0 \\ \frac{f_4}{f_6} & p_8 & 1 \end{pmatrix} \text{ and } \mathbf{P}' = \begin{pmatrix} p_1' & p_2' & 0 \\ 0 & p_5' & 0 \\ \frac{f_2}{f_8} & p_8' & 1 \end{pmatrix}. \quad (7)$$

Although these conditions ensure that the epipolar lines of $\mathcal{I}$ and $\mathcal{I}'$ are parallel, corresponding epipolar lines may lie on different scan-lines. Thus, additional constraints are needed (see figure 3(d)). When $\mathbf{P}$ and $\mathbf{P}'$ map corresponding epipolar lines to the same scan-line the fundamental matrix will be of the form

$$[\mathbf{i}]_\times = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad (8)$$

and therefore the following must hold, $\hat{\mathbf{F}} = \mathbf{P}'^T [\mathbf{i}]_\times \mathbf{P}$. These equations contain two independent constraints, $-f_8 p_5' = f_6 p_5$ and $p_8' f_8 + p_8 f_6 = f_5$, that will be satisfied when

$$\mathbf{P} = \begin{pmatrix} p_1 & p_2 & 0 \\ 0 & p_5 & 0 \\ f_4 & p_8 & f_6 \end{pmatrix} \text{ and } \mathbf{P}' = \begin{pmatrix} p_1' & p_2' & 0 \\ 0 & -p_5 & 0 \\ f_2 & f_5 - p_8 & f_8 \end{pmatrix}. \quad (9)$$

As can be seen from (9), there are 6 free parameters $(p_1, p_2, p_5, p_8, p_1', p_2')$ in the family of perspective transformations that rectify a given stereo image. The parameter p_8 varies with the perspective distortion along the y-axis while p_5 controls scale along the y-axis. Once rectified, image points can be freely moved along the scan-lines and therefore skew and scaling in the x-direction can be applied to both images using p_1, p_2, p_1' and p_2'. The amount of oversampling and under-sampling will depend on the values of these parameters. However, not all of these parameters need to be considered. As discussed earlier, a change of aspect ratio or skew does not alter the local area. But due to the finite size of pixels, changing the aspect ratio or introducing skew will cause resampling effects. Therefore, we impose the following additional constraints:

- Restrict the skew of both images to zero by setting $p_2 = {p_2}' = 0$.
- Ensure the aspect ratio of $\mathcal{I}$ and $\mathcal{I}'$ does not change by requiring $p_5 = p_1$ and $p_1' = p_5'$.

Now, we are left with only two free parameters p_1 and p_8. That is:

$$\mathbf{P} = \begin{pmatrix} p_1 & 0 & 0 \\ 0 & p_1 & 0 \\ f_4 & p_8 & f_6 \end{pmatrix} \text{ and } \mathbf{P}' = \begin{pmatrix} -p_1 & 0 & 0 \\ 0 & -p_1 & 0 \\ f_2 & f_5 - p_8 & f_8 \end{pmatrix}. \quad (10)$$

Since rectification is achieved by projecting the images onto a plane parallel to the line connecting the centers of projection of the two views, we can interpret p_8 as parameterizing the angle of the plane about this line and p_1 as parameterizing the distance. In the image, p_8 controls the amount of perspective distortion along the y-axis and p_1 alters the scale.

4 Finding the Best Transform

From equation (4), the error metric when the transformations $\mathbf{P}$ and $\mathbf{P}'$ are applied to the stereo images is

$$\begin{aligned} \varepsilon \;=\; & \int_{-\frac{h}{2}}^{\frac{h}{2}} \int_{-\frac{w}{2}}^{\frac{w}{2}} \left(\det \frac{\partial \mathbf{P}(\hat{x},\hat{y})}{\partial \mathbf{P}(x,y)} - 1 \right)^2 dx\, dy \; + \\ & \int_{-\frac{h}{2}}^{\frac{h}{2}} \int_{-\frac{w}{2}}^{\frac{w}{2}} \left(\det \frac{\partial \mathbf{P}'(\hat{x},\hat{y})}{\partial \mathbf{P}'(x,y)} - 1 \right)^2 dx\, dy, \end{aligned} \quad (11)$$

where w and h are the width and height of the images, respectively. Our goal is to find the values for the scale parameter p_1 and the perspective parameter p_8 that minimize this objective function.

This integral can be solved in closed form, and we have found that the resulting expression is a 16^{th} degree rational polynomial in p_8. However, the expression is quadratic in the scale parameter. Therefore, given a value for p_8 the value for p_1 that minimizes the integral can be found by solving $\frac{\partial \varepsilon}{\partial p_1} = 0$ to get:

$$p_1^2 = \frac{\iint \frac{dx\,dy}{(f_4 x + p_8 y + f_6)^3} + \iint \frac{dx\,dy}{(f_2 x + (f_5 - p_8) y + f_8)^3}}{\iint \frac{dx\,dy}{(f_4 x + p_8 y + f_6)^6} + \iint \frac{dx\,dy}{(f_2 x + (f_5 - p_8) y + f_8)^6}}. \quad (12)$$

Care must be taken if one of the epipoles falls within the image, because the rectified image will be of infinite extent. In this case, the solution to (12) is $p_1 = 0$, because an image of size zero has a finite loss of pixels whereas an image of infinite size creates an infinite number of pixels. A practical solution can be found by choosing a region of interest to integrate over where the epipoles fall outside the region.

Before solving for the optimal scale parameter, we must obtain a value for the perspective parameter p_8. To find a solution for p_8 we break up the objective function (11) and look at each image separately. For image $\mathcal{I}$, the error metric for re-sampling is

$$\varepsilon_1 = \int_{-\frac{h}{2}}^{\frac{h}{2}} \int_{-\frac{w}{2}}^{\frac{w}{2}} \left(\frac{p_1^2}{(f_4 x + p_8 y + f_6)^3} - 1 \right)^2 dx\, dy, \quad (13)$$

while the error metric for $\mathcal{I}'$ is

$$\varepsilon_2 = \int_{-\frac{h}{2}}^{\frac{h}{2}} \int_{-\frac{w}{2}}^{\frac{w}{2}} \left(\frac{p_1^2}{(f_2 x + (f_5 - p_8) y + f_8)^3} - 1 \right)^2 dx\, dy. \quad (14)$$

After eliminating p_1, the error metric for each image can be rewritten as a function of the location of the epipole for that image and p_8. These functions are symmetric about their

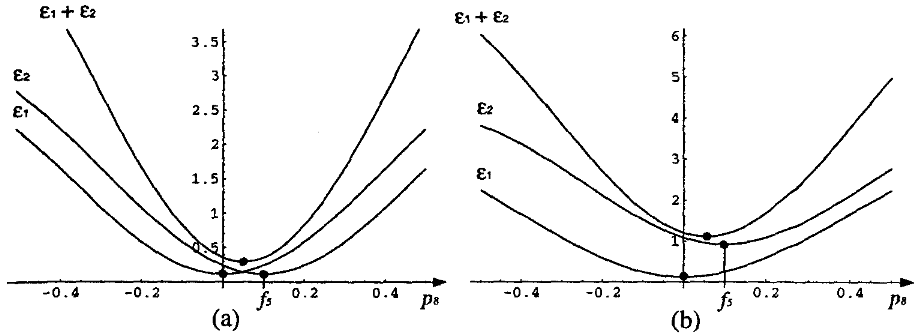


Figure 4: To find an approximate solution for the parameter p_8 we look at the re-sampling metrics ε_1 and ε_2. While these functions are symmetric about their minima, $p_8 = 0$ for ε_1 and $p_8 = f_5$ for ε_2, their exact shapes depend on the locations of the epipoles. Plot (a) shows ε_1 and ε_2 when the epipoles are both located at the same distance from the image centers. In this case the functions are identical except for a shift and therefore the minimum of $\varepsilon_1 + \varepsilon_2$ is $p_8 = \frac{f_5}{2}$. Plot (b) shows ε_1 with an epipole at a distance that is 10 times the width of the image and ε_2 with an epipole at a distance of 3 times. Note that, despite the difference in the locations of the epipoles, the minimum of $\varepsilon_1 + \varepsilon_2$ is still close to $p_8 = \frac{f_5}{2}$.

minima, $p_8 = 0$ for ε_1 and $p_8 = f_5$ for ε_2, and their forms vary with the locations of the epipoles. Figure 4 shows some examples of these functions. When the epipoles are at equal distance from the centers of the images, the shape of each function is identical. Thus, $p_8 = \frac{f_5}{2}$ is the solution that minimizes $\varepsilon_1 + \varepsilon_2$. In general, the solution for p_8 that minimizes $\varepsilon_1 + \varepsilon_2$ lies somewhere between 0 and f_5. We can see this by considering what happens to p_8 as the epipoles move apart. If $\mathbf{e}'$ is fixed at the edge of the image then, in the limit, as $\mathbf{e}$ approaches infinity, the solution is $p_8 = 0$. Likewise, when $\mathbf{e}$ is fixed at the edge of the image then $p_8 = f_5$ in the limit as $\mathbf{e}'$ approaches infinity.

Note that $\varepsilon_1 + \varepsilon_2$ is a convex function in the range 0 to $\frac{f_5}{2}$. Therefore, an optimal solution can be found using $p_8 = \frac{f_5}{2}$ as an initial estimate and applying a simple iterative technique such as gradient descent. During minimization, the scale parameter p_1 is removed from (11) by solving equation (12) using the current estimate for p_8. In practice, $p_8 = \frac{f_5}{2}$ is very close to the optimal solution for the following reasons:

- The magnitude of f_5 is proportional to the amount of tilt. For practical stereo configurations the tilt between the cameras is small and thus f_5 is closer to zero relative to the other entries of the fundamental matrix.
- Even if there is some tilt, most stereo configurations are also verged so that the epipoles are approximately at the same distance, in which case, $p_8 = \frac{f_5}{2}$ is optimal.
- The optimal value for p_8 deviates from $\frac{f_5}{2}$ slowly unless one epipole is close to the image. Except when wide field of view lenses are used, the epipoles are not usually near the stereo images.

5 Results

In this section, we will show examples of rectification applied to a variety of camera geometries. For each stereo image a 500 × 500 pixel checkerboard pattern is used to emphasize the sampling effects of rectification. To visualize the different epipolar geometries a set of corresponding epipolar lines is marked in each of the images.

Figure 5 demonstrates rectification for two geometries typically found in stereo vision. The first stereo image (Figure 5(a)) is in a verged configuration. A more difficult geometry to rectify due to the difference in scale is shown in Figure 5(b). The difference in scale could result from either a change in focal length or forward motion between the stereo cameras. For these two geometries $f_5 = 0$ and therefore the optimal value for p_8 is zero. In 5(b) either the scale of the right image must be increased or the scale of the left image must be decreased in order to match the epipolar lines. This is one instance where many traditional methods for rectification fail by making an *a-priori* decision to either increase the scale of one image or decrease the other. However, by balancing the loss and creation of pixels, our rectification method slightly increases the scale of the right image and slightly decreases the scale of the left image.

In figure 6(a) a stereo pair that has both tilt between the cameras and is not in a verged configuration is rectified. Note how the epipolar lines in the right image get compressed toward the top of the image as a result of the tilt. In order to match these epipolar lines it is necessary to apply perspective distortion along the y-axis. The tilt and the location of the epipoles are chosen such that the approximate solution to p_8 will not be the optimal one. Even in the presence of tilt the results of the optimal and approximate solution are quite similar.

Figure 6(b) shows a geometry that is chosen to create a large difference between the approximate solution and the optimal one. In this geometry one epipole is close to infinity, the other is close to the image boundary and a large amount of tilt is introduced. For this camera configuration the optimal solution considerably improves the result. Although it is

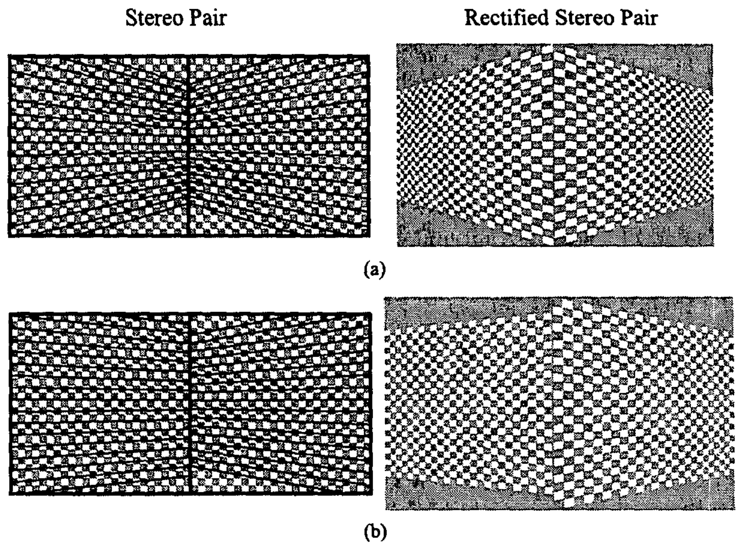


Figure 5: We simulated two epipolar geometries typically found in stereo vision. In each stereo pair a set of epipolar lines is marked, and to visualize the change in local area a checkerboard pattern is used. (a) This pair demonstrates rectification in the presence of a large amount of vergence. (b) The second example shows the rectification results for a difficult geometry. This geometry is difficult because the epipolar lines on the right side are compressed relative to those on the left side, causing a difference in scale between the two views.

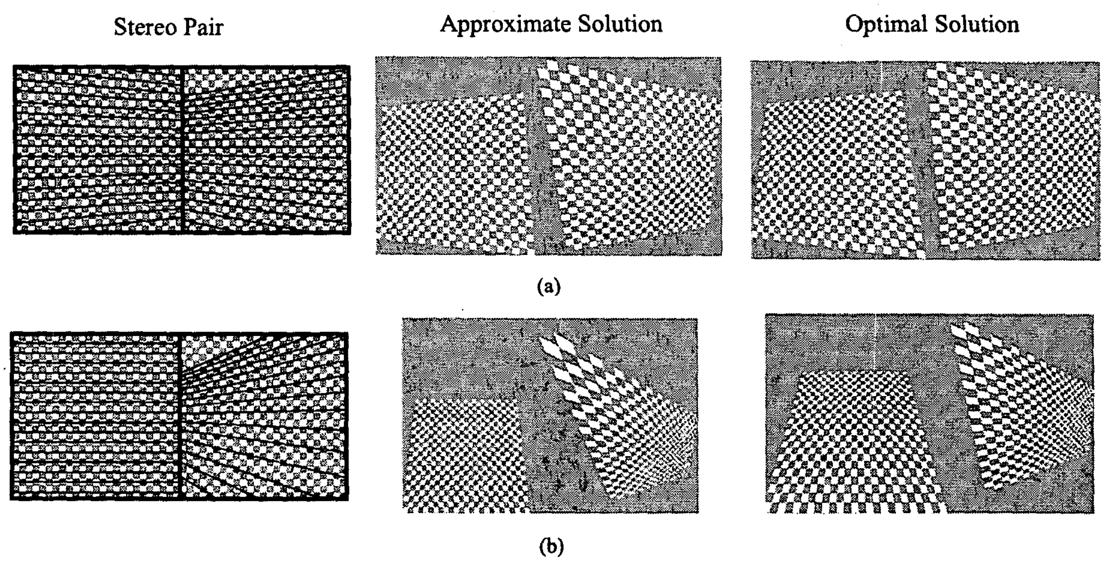


Figure 6: In order to create a situation where the approximate solution $p_8 = \frac{f_6}{2}$ differs from the optimal solution, we simulated epipolar geometries that contain both tilt and asymmetric epipoles. (a) Note that even in the presence of tilt the results are quite similar, demonstrating that for most situations the solution $p_8 = \frac{f_6}{2}$ will suffice. (b) In this example, we chose an epipolar geometry that comes close to maximizing the deviation of the approximate solution from the optimal one. One epipole is close to infinity while the other is located close to the image boundary. In addition a considerable amount of tilt is added. In this situation the optimal solution considerably improves the result.

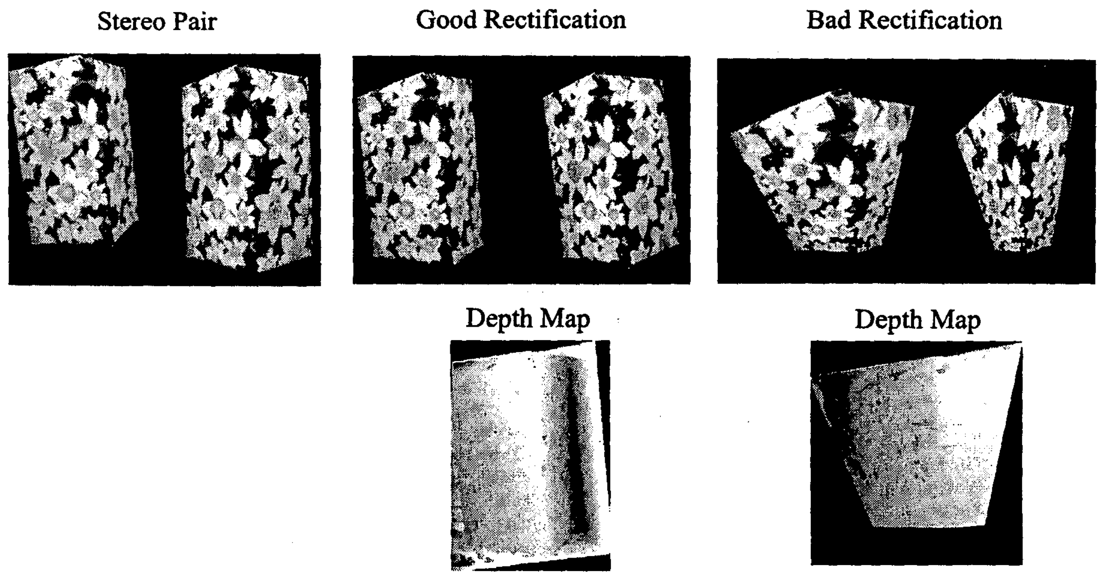


Figure 7: The effect of bad rectification on stereo matching. The first rectified image is the one that minimizes the amount of resampling. The second image is also rectified but is clearly not optimal. On the bottom are the depth maps computed from each of the rectified images. Note that the number of gross mismatches is less in the first case.

unlikely that such a geometry is used in stereo vision, camera configurations like this may be encountered in structure from motion.

Figure 7 shows the effect of a bad rectification on stereo matching. Stereo matching is performed by searching along the scan-lines for corresponding image points that minimize the sum of squared distances between a window of size 7×7 pixels around the image points. Note that the number of gross mismatches is greater in the sub-optimal rectified image.

6 Conclusions

We have presented a method for rectification that determines the perspective transforms that rectify a stereo image pair while preserving the sampling of the original images. First, we proposed using the change in local area to model the amount of pixels lost and created. Then, we derived a parameterization of the class of perspective transforms that rectify a given stereo pair from knowledge of the epipolar geometry. We showed that only two parameters of this class, scale p_1 and perspective distortion p_8, need to be considered when computing the transform that minimizes the change in local area. Lastly, we discussed computation of the optimal values for p_1 and p_8. For most stereo configurations, the approximate solution for p_8 is sufficient. However, for some camera geometries, such as those encountered when recovering structure from motion, computing the optimal value is preferable.

References

[1] V. Arnold. *Mathematical Methods of Classical Mechanics.* Springer-Verlag, 1989.

[2] N. Ayache and C. Hansen. Rectification of images for binocular and trinocular stereovision. In *Proc. Int'l Conf. on Pattern Recognition*, 1988.

[3] O. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint.* MIT press, 1993.

[4] R. Hartley. Theory and practice of projective rectification. *Int'l Journal of Computer Vision*, 35(2), 1999.

[5] C. Loop and Z. Zhang. Computing rectifying homographies for stereo vision. In *Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition*, 2000.

[6] D. Papadimitriou and T. Dennis. Epipolar line estimation and rectification for stereo image pairs. *IEEE Trans. Image Processing*, 5(4):672–676, 1996.

[7] M. Pollefeys, R. Koch, and L. Gool. A simple and efficient rectification method for general motion. In *Proceedings of the 7th International Conference on Computer Vision*, 1999.

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

[9] S. Roy, J. Meunier, and I. Cox. Cylindrical rectification to minimize epipolar distortion. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition*, 1997.

[10] Z. Zhang, R. Deriche, O. Faugeras, and Q. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence Journal*, 78, 1995.