Geometric and Photometric Correction of Projected Rectangular Pictures

Xiang Wang, Reinhard Klette and Bodo Rosenhahn

CITR, Department of Computer Science
The University of Auckland, Auckland, New Zealand

Abstract

This article presents a method for removing geometric distortions in taken images of rectangular planar patterns (such as maps, pictures, or posters) in the absence of any calibration information or explicit knowledge of the imaging device. Our approach removes lens distortion form single-view images. The method is based on Least-Square-Error (LSE) approximations and Hough transforms. We approximate the lens distortion by considering only lower-order terms of the lens distortion. We further apply plane homography to undistort the perspective view and use texture fusion of different views to remove flash lights. Our algorithm transforms images of a planar pattern into an appealing image by eliminating different types of distortion (lens, perspective, flash light).

Keywords: plane homography transform, distortion correction, lens distortion, line detection, Hough transform, flash light

1 Introduction

Lens distortion is still a problem for today's cameras. Projective distortion is always apparent. Flash lights also produce highlights and reflections. This contribution reports about a program which allows the user to invert lens distortion under the assumption that we capture rectangular planar patterns, to remove projective distortion, and to remove flash lights from the image by fusing data of multiple images.

The basic camera model for image formation is the pinhole camera [10]. It assumes that each image point is generated as a direct projection of a 3D point through the (ideal) optical center. But this is just an approximation, since straight lines become curved lines (besides a projection area close to the optical center). Optical physics studies the factors which contribute to different types of lens distortion, which are radial or tangential distortion. Tangential distortion can be neglected [5, 18], and (typically) only lower-order terms $\kappa_1, \kappa_2, \ldots$ of radial distortion must be taken into account [14].

This paper proposes a simple but reliable method to estimate parameters for the correction of radial distortion in an automatic or semiautomatic way. The method can be applied to single images, or can be an initial step in a computer vision process. We tested the algorithm for different cameras or input situations. We also show that an undistorted image (showing an ideal rectangle) can be back projected into an ideal projective image, see Figure 1. Based on our distortion correction method, we apply image mosaics [16] and merging [1] methods to remove effects of flash lights.

Figure 1: Distorted input image (left), and image obtained under ideal central projection (right).

2 Model of 2D Distortion

Figure 1 demonstrates lens and projective distortion: on the left both is apparent, and on the right only projective distortion. Lens distortion can be removed by camera calibration [18]. We do no require camera calibration, we only analyze a single image showing a rectangular object. The removal of lens distortion in the absence of any calibration information or explicit knowledge of the imaging device is the main subject in this paper.

2.1 Single parameter for lens distortion

The mapping between 3D scene and 2D image points can be decomposed into a perspective projection and a function that models the deviation from the ideal pinhole camera [18].

The pinhole camera-centered coordinate system projects a 3D scene point $P = (X, Y, Z)$ onto an image point $p = (x, y)$ in the image plane:

$$x = f\frac{X}{Z}, \quad \text{and} \quad y = f\frac{Y}{Z} \tag{1}$$

where f denotes the (actual) focal length.

A general model of lens distortion has been discussed in [15]. Give a ”distorted” image point $p_d = (x_d, y_d)$, we can obtain the ”undistorted” image point $p_u = (x_u, y_u)$ as follows [4]:

$$x_u = c_x + (x_d - c_x)(1 + \kappa_1 r_d^2 + \kappa_2 r_d^4 + \cdots) + p_1[2(x_d - c_x) + r_d^2] + 2p_2(x_d - c_x)(y_d - c_y) \tag{2}$$

$$y_u = c_y + (y_d - c_y)(1 + \kappa_1 r_d^2 + \kappa_2 r_d^4 + \cdots) + p_2[2(y_d - c_y) + r_d^2] + 2p_1(x_d - c_x)(y_d - c_y) \tag{3}$$

where (c_x, c_y) are the coordinates of the center of distortion and $r_d = \sqrt{(x_d - c_x)^2 + (y_d - c_y)^2}$.

We can disregard the tangential components:

$$\begin{aligned} x_u &= c_x + (x_d - c_x)(1 + \kappa_1 r_d^2 + \kappa_2 r_d^4 + \cdots) \\ y_u &= c_y + (y_d - c_y)(1 + \kappa_1 r_d^2 + \kappa_2 r_d^4 + \cdots) \end{aligned} \tag{4}$$

The lens distortion model is often specified by using infinite series. However, there is experimental evidence [7, 14] that approximating these series with only the lower-order components corrects more than 90% of the radial distortion. With using only the first-order radial symmetric distortion parameter κ_1 we can achieve a precision of about 0.1 pixels in the image space (using today's lenses).

Therefore, Equation (4) can be approximated as:

$$\begin{aligned} x_u &= x_d + (x_d - c_x)(\kappa_1 r_d^2) \\ y_u &= y_d + (y_d - c_y)(\kappa_1 r_d^2) \end{aligned} \tag{5}$$

Given distorted point coordinates, we calculate undistorted point coordinates by using only one coefficient κ_1, where (c_x, c_y) are the coordinates of the center of distortion. Next we deal with the inversion of Equation 5 for modeling the distorted image, since we only have the given digital image (without any additional calibration information or explicit knowledge of the imaging device) for estimating the κ_1 value of the given digital image.

2.2 Distorted coordinates

We use the first-order radial symmetric distortion parameter κ_1, so the undistorted coordinates are given as follows:

$$\begin{aligned} x_u &= x_d + (x_d - c_x)(\kappa_1 r_d^2) \\ y_u &= y_d + (y_d - c_y)(\kappa_1 r_d^2) \end{aligned}$$

where r_d is as defined above. The coordinates of the *center of distortion* are (c_x, c_y).

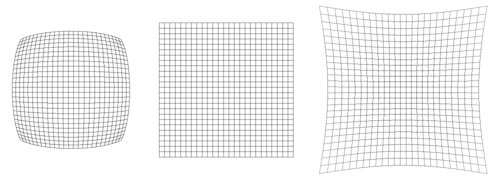

Figure 2: Left: barrel distortion for $\kappa > 0$. Middle: ideal (undistorted) image for $\kappa = 0$. Right: pincushion distortion for $\kappa < 0$.

The distorted coordinates can be expressed as a function of the undistorted coordinates; they are given as a solution of the following equation:

$$r_u = r_d(1 + k_1 r_d^2)$$

with $r_u = \sqrt{{x_u}^2 + y_u^2}$.

There are several papers [4, 5, 6, 13, 19] on the relationship between r_d and r_u, which is a polynomial of degree three in r_d of the form

$$r_d^3 + c r_d + d = 0$$

with $c = 1/\kappa_1$ and $d = -c r_u$. It can be solved using the Cardan method, which is a direct method for solving polynomials of degree three. This method provides the only real solution of the polynomial if κ_1 is positive. This solution is as follows:

$$r_d \doteq r_u(1 + \kappa_1 r_u^2)$$

The distorted coordinates are then given by the following:

$$\begin{aligned} x_d &= c_x + (x_u - c_x)\frac{r_d}{r_u} \\ y_d &= c_y + (y_u - c_y)\frac{r_d}{r_u} \end{aligned}$$

Figure 2 shows the resulting images of a rectilinear grid for three different values of κ_1.

3 Impacts of Set-Up Parameters

This section discusses the influence of different parameters on image distortion; based on this we select the ones which are most crucial for removing distortions.

3.1 Distance to object plane

[17] discusses a wide-angle camera; assuming fixed focal length, the size of the image plane is always constant. No matter how the distance varies between object and camera, the “curvature” and the location of each pixel in the distorted image array remain constant within the image plane. In

Figure 3 we assume that we have an object plane (containing a drawing, a map, and so forth) and the image plane, and that both planes are parallel. O represents the (focal point of the) lens, and D_1, D_2 are two different distances measured along the central projection ray (i.e., the optical axis) between lens and object plane. We assume that D_1, D_2 are possible distances for the given focal length. The segment OO' corresponds to this effective focal length, and the size of the image plane remains constant due to having a constant focal length.

Two corresponding lines L_2 and L_1 are assumed to be in the further-away or closer object plane, respectively, and they are both mapped into the same distorted line C_2 in the image plane. The difference is within the range of the depth of field; when the object is, of course, that we capture a smaller area if the object plane moves closer to the camera.

It follows that we are free to resize (i.e., scale) the input image for better accuracy of results, or for reduction of computing time; resizing will not effect the geometric appearance of the lens distortion in the image.

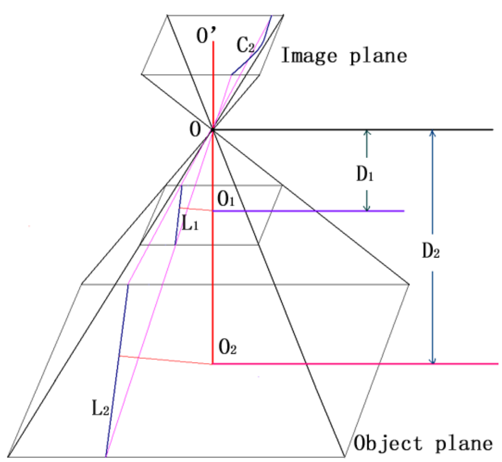

Figure 3: A sketch of geometric relations between image plane, lens and object plane.

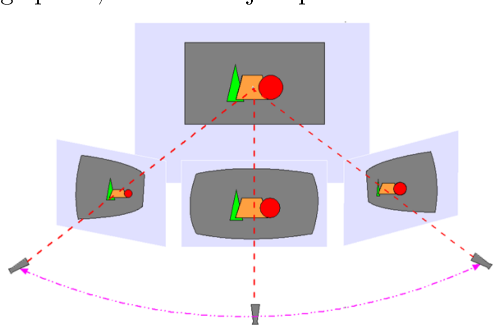

Figure 4: Varying angles of camera views.

3.2 Distortion center

Under ideal projection conditions, the lens distortion center is the center of the given image. But practical experience shows that we do not deal with such an ideal case normally. [20] presents a taxonomy that includes 15 different definitions of image centers, which move in the image plane as lens parameters change. The authors conclude that the accuracy of the image center is critical for the accuracy of camera calibration, and that errors in estimating locations of image centers (following different center definitions and different lens settings) makes the calibration problem especially hard to solve. Lens distortion centers move with varying imaging set-ups (e.g., changes of focal length, different distances to the object, different digital image settings, different angles, and so forth).

We use a least-square error approach and Hough transform for line fitting (i.e., these are two alternative methods which also can be combined), and this will allow to estimate center and coefficients of lens distortion within one procedure, if the distortion center is relatively close to the image's center (say, in about 10% pixel distance at most compared to the image's size).

3.3 Viewing angle

Figure 4 illustrates the projective situation for image and object plane if the viewing angle of the camera changes. The originally rectangular object (in the object plane) becomes (under ideal projective conditions) a trapeziform object (possibly with no parallel edges at all) in the image plane. However, the lens distortion will also add curvatures to these edges.

We use a plane homography method to transform the ("curved") trapeziform object back into a ("curved") rectangular object. As a second step we then correct the lens distortion.

3.4 Focal length

[12] discusses that a change in focal length will also change the coefficient(s) of lens distortion. However, in our case we do not have to calibrate for multiple images taken by the same camera, and dealing with a single image only allows to neglect focal length. We only use a method that relies on the assumption that pure radial distortion transforms straight lines into curved lines. [Note: for combining, e.g. via stitching, a small number of different images of the same object, we still consider this as a reasonable approach because these images should then be taken all with about the same focal length.]

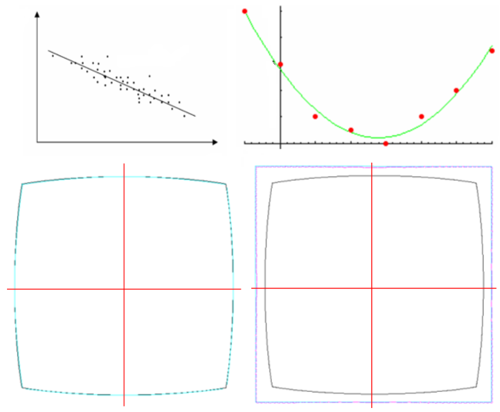

Figure 5: Upper left: least-square line fitting. Upper right: least-square curve matching based on varying κ_1. Lower left: all four calculated curved edges of one rectangular object. Lower right: the final mapping into a rectangle.

4 The Method

Our method for correcting lens and projective distortion consists of several steps. First, the user has to identify the four corners $((x_1, y_1), (x_2, y_2), (x_3, y_3)$, and $(x_4, y_4))$ of the pictured rectangular object. Second, we do plane homography to transform these four object corners into a general position such that they are the corners of a rectangular object (at some distance) in an object plane parallel to the image plane. Third, we do edge detection (between the identified corners) to localize the ("curved") edges of the rectangular object; this is followed by a least-square error match of a second order curve. Finally, the following two methods can be used alternatively (or combined) to do the distortion correction. Possibly, we are also interested in back projecting the rectangular undistorted object into the original object plane.

4.1 Edge detection and curve matching

The use of a single edge detection method [2, 9] proved to be insufficient in general. Typically, we combined first a first order (derivatives) edge detector such as Sobel, with a second order edge detector (such as LoG); then we used the Canny operator, and final results were derived from both.

Following standard least-square error (LSE) polynomial curve fitting, we fit $y = a + bx + cx^2$ as follows:

$$\sum_{i=1}^{n} y_i = a\sum_{i=1}^{n} 1 + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} x_i^2$$
$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i^3$$
$$\sum_{i=1}^{n} x_i^2 y_i = a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^4$$

We obtain a, b and c by solve the given LSE problem defined by these polynomial equations. The solution is used to match the detected edges (see upper right and lower left in Figure 5).

Similarly, we use LSE line fitting; the equation $y = a + bx$ expands into

$$\sum_{i=1}^{n} y_i = a\sum_{i=1}^{n} 1 + b\sum_{i=1}^{n} x_i$$
$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2$$

The LSE of the resulting line is as follows:

$$\prod = \sum_{i=1}^{n}[y_i - f_{x_i}]^2 = \sum_{i=1}^{n}[y_i - a + bx_i]^2$$

and this is used for distortion correction. The minimum error, which defines the κ_1 value of distortion, should map all four curved lines into straight lines, if properly calculated. [Note: improper calculations might be due, e.g., by a large distance between distortion center and image center.]

4.2 Plane homography

We assume four corner points (x_1, y_1), (x_2, y_2), (x_3, y_3), and (x_4, y_4) of the rectangular object. We use a plane homography method to project those four points into a rectangle having the corners (S_1, T_1), (S_2, T_2), (S_3, T_3), and (S_4, T_4) (see Figure 6). This transform is given by the solution of the following equation:

$$A * T = B \qquad (6)$$

by A given as

$$\begin{pmatrix} -x_1 & -y_1 & -1 & 0 & 0 & 0 & S_1x_1 & S_1y_1 \\ 0 & 0 & 0 & -x_1 & -y_1 & -1 & T_1x_1 & T_1y_1 \\ -x_2 & -y_2 & -1 & 0 & 0 & 0 & S_2x_2 & S_2y_2 \\ 0 & 0 & 0 & -x_2 & -y_2 & -1 & T_2x_2 & T_2y_2 \\ -x_3 & -y_3 & -1 & 0 & 0 & 0 & S_3x_3 & S_3y_3 \\ 0 & 0 & 0 & -x_3 & -y_3 & -1 & T_3x_3 & T_3y_3 \\ -x_4 & -y_4 & -1 & 0 & 0 & 0 & S_4x_4 & S_4y_4 \\ 0 & 0 & 0 & -x_4 & -y_4 & -1 & T_4x_4 & T_4y_4 \end{pmatrix}$$

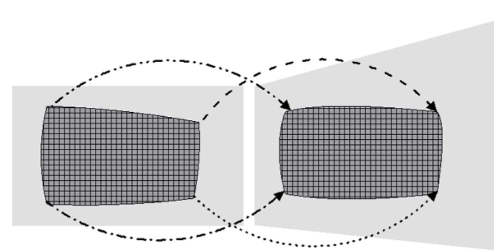

Figure 6: The plane homography transform.

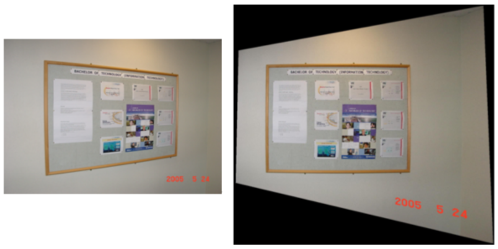

Figure 7: The input image (on the left) is processed by plane homography into the output image (on the right).

and the vectors

$$T = (a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32})^T$$
$$B = (S_1, T_1, S_2, T_2, S_3, T_3, S_4, T_4)^T .$$

The vector T represents eight values of the 3x3 transformation matrix H (see below) which incorporates (in general) rotation, translation, scaling, skewing, and stretching as well as perspective distortion. Since we only consider perspective distortion, H simplifies to

$$H = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

in this case, where $a_{33} = 1$.

Figure 7 shows on the left an image of a rectangular object, and on the right the plane homography transform into the target object which is positioned parallel to the plane image.

4.3 Correction of lens distortion

Finally we use line detection for estimating the value of κ_1 and the center of distortion. We experimented with two common methods for line detection, and we can propose both as alternatives, or for possible combination. The first is LSE line fitting (as described above), and the second is Hough transform.[8]

We generate values in the Hough space for different values of κ_1, one such Hough map (only around the four expected accumulation points!) for each κ_1. Then we apply for find the exact positions of the accumulation points (as illustrated in Figure 8) together with a weight of those points; we just use a 3×3 mask and search for the maximum sum of Hough values around the four expected accumulation points. The Hough map with the maximum weights defines the κ_1 value and the corresponding center of distortion.

5 Experimental Results

Figure 9 shows an example. The image on the upper left is the original input image. The image on the upper right is the image transformed by projective transform and lens distortion. The image on the lower left is the final undistorted image, and the image on the lower right is the back projection of this undistorted image into the object plane (showing the original image as it appears without lens distortion).

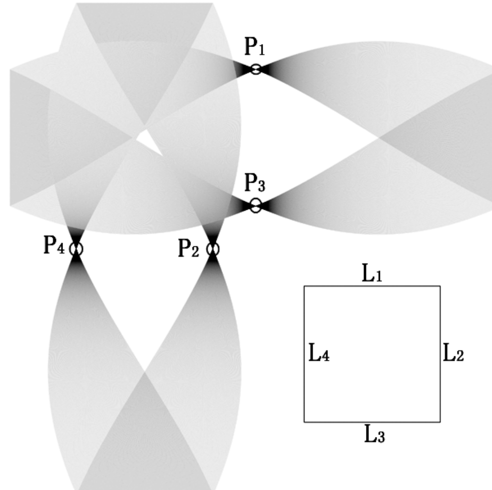

Figure 8: The shown four accumulation points in the Hough space define four edges of the rectangle.

Figure 10 illustrates our approach to remove flash lights. The top row shows two input images. The images in the middle are geometrically undistorted, where the homography transform was applied to both images individually. Then we use the Harris corner detector [11] to find approximate locations of a flash light in each image. Finally, we using an image registration [3] and mosaicing method [16] to merge both images to replace the flash areas. An image merging [1] method is used to adjust intensity levels (the image at the bottom of Figure 10 is final result).

6 Conclusion and Summary

This contribution presents a method for lens undistortion and projective correction of images showing rectangular objects. The method can be applied to images taken under different angles, using arbitrary focal length, position and image size, as long as a rectangular plane object is captured, and the center of distortion is still "close" to the center of the image. Flash removal can be done based on this method. This might be an important initial step for some of the computer vision applications. It is valuable under circumstances where we do not know any calibration information about the camera at all.

Figure 9: Upper left: the original (distorted) image. Upper right: image processed by plane homography. Lower left: undistorted image. Lower right: back transform of image by plane homography.

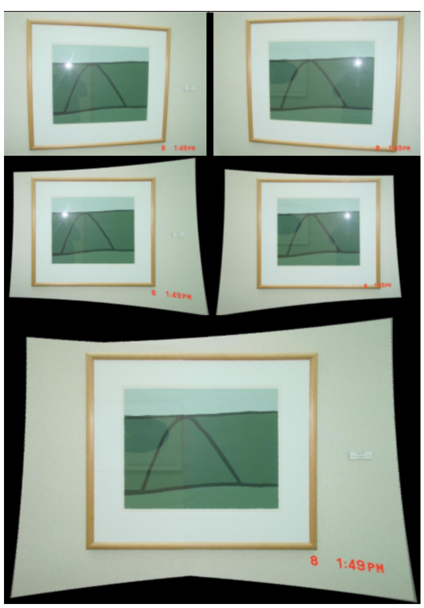

Figure 10: Flash removal with two images.

References

[1] P. J. Burt and E. H. Adelson. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics*, **2(4)**:217C236, 1983.

[2] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis Machine Intelligence*, **8**:679–697, 1986.

[3] C. Y. Chen and R. Klette. Image stitching - comparisons and new techniques. *IVCNZ-Imaging Vision Computing New Zealand* ,101-105, 1998.

[4] R. Cucchiara, C. Grana, A. Prati, and R. Vezzani. A Hough transform-based method for radial lens distortion correction. In Proc. Int. IEEE Conf. *Image Analysis Processing*, pages 182-187, 2003.

[5] F. Devernay and O. Faugeras. Straight lines have to be straight. Automatic calibration and removal of distortion from scenes of structured environments. *Machine Vision and Applications*, **13**:14–24, 2001.

[6] H. Farid and A. C. Popescu. Blind removal of lens distortion. In Proc. *ICCV-International Conference on Computer Vision*, pages 76–81. 2001.

[7] O. Faugeras and G. Toscani. Structure from motion using the reconstruction and reprojection technique. In Proc. IEEE Workshop *Computer Vision*, pages 345–348, 1987.

[8] J. Illingworth and J. Kittler. A survey of the Hough transform. *Computer Vision Graphics Image Processing*, **43**:221–238, 1988.

[9] R. Klette and P. Zamperoni. *Handbook of Image Processing Operators.* Wiley & Sons, Chichester, 1996.

[10] R. Klette, K. Schlüns, and A. Koschan. *Computer Vision - Three-Dimensional Data from Digital Images.* Springer, Singapore, 1998.

[11] C. Harris and M. Stephens. A combined corner and edge detector. *Proc. 4th Alvey Vision Conference*, 189-192, 1988.

[12] Y. C. Lin and C. S. Fuh. Correcting distortion for digital cameras. *Proc. Natl. Sci. Counc. ROC (A)*, **24**:115–119. 2000.

[13] L. Ma, Y. Q. Chen, and K. L. Moore. Blind detection and compensation of camera lens geometric distortions. *CoRR, cs.CV/0405095, http://arxiv.org/abs/cs.CV/0405095*, 2004.

[14] R. Mohr and B. Triggs. Projective Geometry for Image Analysis " In Proc. *ISPRS-International Society for Photogrammetry and Remote Sensing*, tutorial, 1996.

[15] C. C. Slama. *Manual of Photogrammetry*, Fourth Edition. American Society of Photogrammetry and Remote Sensing, Falls Church, 1980.

[16] R. Szeliski and H. Y. Shum. Creating full view panoramic image mosaic and environment maps. *In Proc. SIGGRAPH97* , 251-258 1997.

[17] H. Tian, T. Srikanthan, V. K. Asari, and S. K. Lam. Study on the effect of object to camera distance on polynomial expansion coefficients in barrel distortion correction. In Proc. *SSIAI*, pages 255–259. 2002.

[18] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE J. Robotics Automation*, **3**:323–344, 1987.

[19] R. J. Valkenburg and P. L. Evans. Lens distortion calibration by straightening lines. In Proc. *Imaging Vision Computing New Zealand*, pages 65–69, 2002.

[20] R. G. Wilson and S. A. Shafer. What is the center of the image?. *J. Optical Society of America A*, **11**:2946–2955, 1994.