8-1-2006

Metric Measurements on a Plane from a Single Image

Micah K. Johnson
Dartmouth College

Hany Farid
Dartmouth College

Metric Measurements on a Plane from a Single Image

Micah K. Johnson and Hany Farid
Department of Computer Science
Dartmouth College
Hanover NH 03755

Abstract

The past decade has seen considerable advances in the application of principles from projective geometry to problems in image analysis and computer vision. In this paper, we review a subset of this work, and leverage these results for the purpose of forensic analysis. Specifically, we review three techniques for making metric measurements on planar surfaces from a single image. The resulting techniques should prove useful in forensic settings where real-world measurements are required.

Figure 1: Shown on the left is the original image. Shown on the top right is a close-up of the license plate which is largely illegible. Shown in the bottom right is the result of planar rectification followed by histogram equalization.

1 Introduction

Shown in Figure 1 is an image of a license plate that is largely illegible. Also shown in this figure (bottom right panel) is the result of transforming the license plate as if it were viewed head-on. This rectified image clearly reveals the license plate number. The mathematics that allows for this sort of planar rectification is fairly well understood. Here, we review several tools from projective geometry that allow for the rectification of planar surfaces and, under certain conditions, the ability to make real-world measurements from a planar surface.

We survey three techniques for the rectification of planar surfaces imaged under perspective projection. Each method requires only a single image. The first, and most direct, method exploits knowledge of polygons of known shape (e.g., street sign, license plate, lettering on a billboard). The second method requires knowledge of two or more vanishing points on a plane and, for example, a pair of known angles on the plane. The third method requires two or more coplanar circles (e.g., car wheels). In each case, we recover the world to image transformation of the planar surface, thereby allowing metric measurements to be made on the plane.

2 Methods

2.1 Polygons

As described in [3], we review how known 2-D shapes (e.g., squares, rectangles, letters, etc.) on a planar surface in the world can be used to make metric measurements on the plane.

Under an ideal pinhole camera model, points on a plane, $\vec{X}$, in the world coordinate system are imaged to the image plane with coordinates $\vec{x}$, given by:

$$\vec{x} \;=\; H\vec{X}, \tag{1}$$

where both points are homogeneous 3-vectors in their respective coordinate systems. Note that under this model, straight lines in the world are imaged to straight lines in the image. Under a more realistic lens model, however, straight lines in the world are distorted and, therefore, not imaged as straight lines. Since the technique described below requires the extraction of straight lines, such lens distortion needs to be removed, Appendix A.

It is easy to see[1] that:

$$\begin{aligned}
\vec{x} \times H\vec{X} &= \vec{0} \\
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} h_1X_1 + h_2X_2 + h_3X_3 \\ h_4X_1 + h_5X_2 + h_6X_3 \\ h_7X_1 + h_8X_2 + h_9X_3 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned} \tag{2}$$

where $\times$ denotes a cross product. Evaluating the cross product on the left-hand side yields:

$$\begin{pmatrix} x_2(h_7X_1 + h_8X_2 + h_9X_3) - x_3(h_4X_1 + h_5X_2 + h_6X_3) \\ x_3(h_1X_1 + h_2X_2 + h_3X_3) - x_1(h_7X_1 + h_8X_2 + h_9X_3) \\ x_1(h_4X_1 + h_5X_2 + h_6X_3) - x_2(h_1X_1 + h_2X_2 + h_3X_3) \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3}$$

Note that this constraint is linear in h_i, the coefficients of H. Re-ordering the terms yields the following system of linear equations for the components of the matrix H:

$$\begin{pmatrix} 0 & 0 & 0 & -x_3X_1 & -x_3X_2 & -x_3X_3 & x_2X_1 & x_2X_2 & x_2X_3 \\ x_3X_1 & x_3X_2 & x_3X_3 & 0 & 0 & 0 & -x_1X_1 & -x_1X_2 & -x_1X_3 \\ -x_2X_1 & -x_2X_2 & -x_2X_3 & x_1X_1 & x_1X_2 & x_1X_3 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$M\vec{h} \;=\; \vec{0}. \tag{4}$$

Given the known coordinates of a point, $\vec{X}$, on a plane in the world and its corresponding projected coordinates, $\vec{x}$, the above system seemingly provides three constraints in the nine unknowns of $\vec{h}$.

[1] The cross product is defined as: $\vec{a} \times \vec{b} = \vec{n}\|a\|\|b\| \sin(\theta)$, where $\vec{n}$ is orthogonal to $\vec{a}$ and $\vec{b}$, and θ is the angle between $\vec{a}$ and $\vec{b}$. Note that if $\vec{a} = \vec{b}$, then $\theta = 0$ and $\vec{a} \times \vec{b} = 0$.

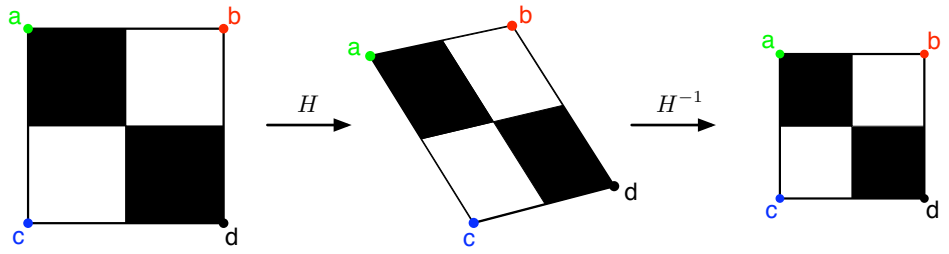

Figure 2: The transformation H from world to image coordinates and the transformation H^{-1} from image to world coordinates. Knowledge of the transformation H^{-1} and the appropriate scale factor allows for metric measurements to be made on the plane $abcd$.

Note, however, that the rows of the matrix M are not linearly independent (the third row is a linear combination of the first two rows). As such, this system provides only two constraints in the nine unknowns. In order to solve for the projective transformation matrix H, four or more points[2] with known coordinates $\vec{X}$ and $\vec{x}$ are packed into the rows of the matrix M. The vector $\vec{h}$ that satisfies $\min_{\vec{h}} \|M\vec{h}\|$, such that $\|\vec{h}\|^2 = 1$, is the minimal eigenvalue eigenvector of $M^T M$.

The estimation of H is determined up to an unknown scale factor. From a single image, a known length on the world plane is required in order to determine this scale factor. With a known projective transformation matrix H, the image is warped according to H^{-1} to yield a rectified image, from which measurements can be made, Figure 2.

2.2 Vanishing Points

As described in [5], we review how two or more vanishing points can be used to make metric measurements on a planar surface in the world.

Recall that under an ideal pinhole camera model, points, $\vec{X}$, on a plane in the world coordinate system are imaged to the image plane with coordinates $\vec{x}$, given by $\vec{x} = H\vec{X}$. Here we consider the inverse mapping:

$$\begin{aligned} \vec{X} &= H^{-1}\vec{x} \\ &= (SAP)\vec{x}, \end{aligned} \tag{5}$$

where both points are homogeneous 3-vectors in their respective coordinate systems. The projective transformation matrix H^{-1} is uniquely decomposed into a product of three matrices: a similarity matrix S, an affine matrix A and a pure projective matrix P. As in the previous section, lens distortion needs to be removed so that straight lines in the world are imaged to straight lines in the image, Appendix A.

[2]With four points, the rank of matrix M is 8, one less than the number of unknowns in $\vec{h}$. This, however, is sufficient, as we only can recover $\vec{h}$ to within a scale factor.

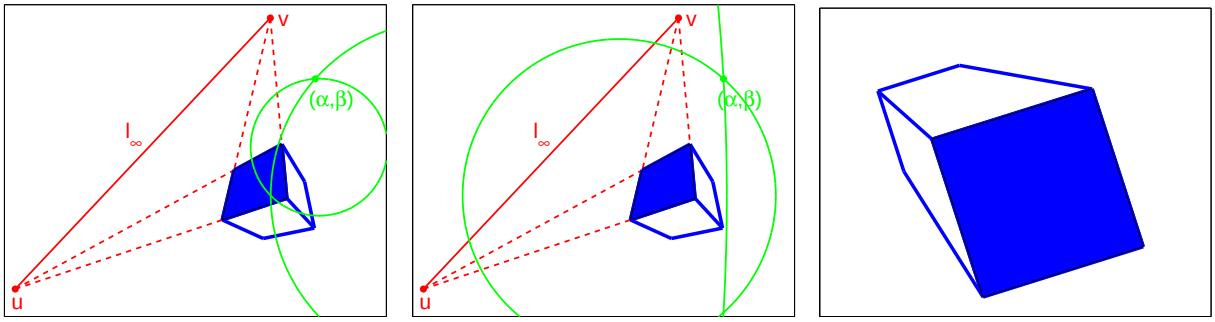

Figure 3: A cube synthesized under perspective projection. Parallel lines on a cube face intersect at vanishing points u and v. The line l_∞ specifies the pure projective matrix, Equation (6). Shown in the left-most panel are the constraint circles from known angles. Shown in the middle panel are the constraint circles from known length ratios. The intersection (α, β) of these constraint circles yields the affine matrix A, Equation (7). Shown in the right-most panel is the rectified top face of the cube (within an arbitrary rotation, translation and scale factor).

The pure projective matrix P is given by:

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{pmatrix}, \tag{6}$$

where $l_\infty = \begin{pmatrix} l_1 & l_2 & l_3 \end{pmatrix}^T$ is the vanishing line of the plane (parameterized as a homogeneous 3-vector: $l_1 x + l_2 y + l_3 = 0$) defined to be the line on which the vanishing points of the plane lie, Figure 3. Shown in this figure is a cube synthesized under perspective projection. Each pair of parallel lines on the top cube face intersect at vanishing points u and v. The line l_∞ is the line that connects these two points.

After applying the pure projective matrix P to the image coordinates $\vec{x}$, the affine matrix A is given by:

$$A = \begin{pmatrix} \frac{1}{\beta} & -\frac{\alpha}{\beta} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{7}$$

where the coefficients α and β are estimated as follows. Given a known angle θ on the world plane between two lines l and m (parameterized as a homogeneous 3-vector), it can be shown that α and β lie on a circle with center:

$$(c_\alpha, c_\beta) = \left(\frac{a+b}{2}, \frac{a-b}{2} \cot(\theta) \right), \tag{8}$$

and with radius:

$$r = \left| \frac{a-b}{2 \sin(\theta)} \right|, \tag{9}$$

where $a = -l_2/l_1$ and $b = -m_2/m_1$.

Alternatively, a known length ratio between two non-parallel line segments l and m (parameterized by a pair of coordinates denoting each segment's end-points) on the world plane also provides a constraint circle with center:

$$(c_\alpha, c_\beta) = \left(\frac{\Delta l_1 \Delta l_2 - \rho^2 \Delta m_1 \Delta m_2}{\Delta l_2^2 - \rho^2 \Delta m_2^2}, 0 \right) \tag{10}$$

and with radius:

$$r = \left| \frac{\rho(\Delta m_1 \Delta l_2 - \Delta l_1 \Delta m_2)}{\Delta l_2^2 - \rho^2 \Delta m_2^2} \right|, \tag{11}$$

where ρ is the known length ratio between l and m, and Δl_1 and Δl_2 are the differences between the first and second coordinates, respectively, of the line segment l. Other constraints that may be employed are equal but unknown angles or five orthogonal line pairs [5].

Two independent constraints of either known angles or known length ratios, or combinations thereof, generate a pair of circles whose intersection is (α, β), Appendix C, yielding the desired matrix A.

The final similarity matrix S is given by:

$$S = \begin{pmatrix} sr_1 & sr_2 & t_x \\ sr_3 & sr_4 & t_y \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} sR & \vec{t} \\ \vec{0}^T & 1 \end{pmatrix}, \tag{12}$$

where s is an isotropic scaling, R is a rotation matrix, and $\vec{t}$ is a translation vector. Only the scale factor s is required in order to make absolute Euclidean measurements on a world plane. From a single image, a known length on the world plane is required in order to determine this scale factor.

The desired projective transformation matrix H^{-1} is computed by simply multiplying the matrices given in Equations (12), (7) and (6). The image is then warped according to H^{-1} to yield a rectified image, from which measurements can be made, Figure 3.

2.3 Circles

Recall that the mapping from image to world coordinates can be decomposed into a product of three matrices, $H^{-1} = SAP$, Equation (5). In the previous section, the pure projective matrix P and the affine matrix A were estimated from vanishing points. In this section, these matrices are estimated from two or more coplanar circles – the intersection of these circles yields the coefficients of P and A.

The equation of a circle centered at the point (x_0, y_0) with radius r is given by:

$$(x - x_0)^2 + (y - y_0)^2 = r^2. \tag{13}$$

In homogeneous coordinates, the equation of a circle can be expressed as:

$$\vec{x}^T C \vec{x} = 0, \tag{14}$$

where $\vec{x} = \begin{pmatrix} x & y & 1 \end{pmatrix}^T$, and the conic coefficient matrix is given by:

$$C = \begin{pmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ -x_0 & -y_0 & x_0^2 + y_0^2 - r^2 \end{pmatrix}. \tag{15}$$

All points, $\vec{x}$, that lie on the circle satisfy Equation (14). The circular points, $\vec{I} = \begin{pmatrix} 1 & i & 0 \end{pmatrix}^T$ and $\vec{J} = \begin{pmatrix} 1 & -i & 0 \end{pmatrix}^T$, with $i = \sqrt{-1}$, lie on every circle in a plane. Note that since the circular points lie on every circle in the plane, any two circles on the plane must intersect at these points.

Consider now the effect of the projective transform H on the circular points $\vec{I}$ and $\vec{J}$. The mapping from image to world coordinates is:

$$H^{-1} = SAP, \tag{16}$$

The mapping from world to image coordinates is:

$$H = (SAP)^{-1} = P^{-1}A^{-1}S^{-1}, \tag{17}$$

where

$$P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -l_1/l_3 & -l_2/l_3 & 1/l_3 \end{pmatrix} \quad \text{and} \quad A^{-1} = \begin{pmatrix} \beta & \alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{18}$$

and where S^{-1} is the similarity matrix, whose inverse is given in Equation (12). The circular points are mapped under this projective transform H:

$$H\vec{I} = P^{-1}A^{-1}S^{-1}\vec{I} \quad \text{and} \quad H\vec{J} = P^{-1}A^{-1}S^{-1}\vec{J}. \tag{19}$$

For our purposes, the similarity matrix S^{-1} can, for now, be ignored because the circular points are either invariant or are swapped under this transformation [3]. Without loss of generality, we can assume that the circular points are invariant to S^{-1}, in which case:

$$H\vec{I} = P^{-1}A^{-1}\vec{I} = P^{-1}\begin{pmatrix} \beta + i\alpha \\ i \\ 0 \end{pmatrix} = \begin{pmatrix} \beta + i\alpha \\ i \\ -l_1/l_3(\beta + i\alpha) - il_2/l_3 \end{pmatrix}. \tag{20}$$

The projection of $\vec{J}$ is the complex conjugate of $H\vec{I}$. Note that $H\vec{I}$ and $H\vec{J}$ contain the required coefficients of the desired matrices P and A. Under the projective transform H, the intersections of any two coplanar circles are preserved. These circles are projected to ellipses in the image. As a result, we seek to compute the intersection of these ellipses, two of which correspond to $H\vec{I}$ and $H\vec{J}$.

Consider the projection of two coplanar circles to ellipses, each of which is given by:

$$d_1x^2 + d_2xy + d_3y^2 + d_4x + d_5y + d_6 = 0 \tag{21}$$

$$e_1x^2 + e_2xy + e_3y^2 + e_4x + e_5y + e_6 = 0, \tag{22}$$

The coefficients of an ellipse can be determined from the manual or automatic extraction of points on the ellipse, Appendix D. The points of intersection, $H\vec{I}$ and $H\vec{J}$, of the two ellipses satisfy both of

these equations. These equations can be expressed as quadratic equations in x with coefficients that are a function of y:

$$f(x,y) = d_1x^2 + (d_2y + d_4)x + (d_3y^2 + d_5y + d_6) \quad (23)$$
$$g(x,y) = e_1x^2 + (e_2y + e_4)x + (e_3y^2 + e_5y + e_6). \quad (24)$$

The resultant $\text{Res}(f,g)$ (see Appendix E for details) is a quartic polynomial in y given by:

$$q(y) = r_4y^4 + r_3y^3 + r_2y^2 + r_1y + r_0. \quad (25)$$

The roots of $q(y)$ yield the y-coordinates of the intersection points, $H\vec{I}$ and $H\vec{J}$ – the corresponding x-coordinates can be determined from Equation (58). Solving analytically for the roots of $q(y)$ is cumbersome. Instead, we solve for the roots numerically by forming the companion matrix for $q(y)$:

$$Q = \begin{pmatrix} 0 & 0 & 0 & -r_0/r_4 \\ 1 & 0 & 0 & -r_1/r_4 \\ 0 & 1 & 0 & -r_2/r_4 \\ 0 & 0 & 1 & -r_3/r_4 \end{pmatrix}. \quad (26)$$

The four eigenvalues of the companion matrix Q are the roots of the polynomial $q(y)$, corresponding to the intersection of the two ellipses. If the ellipses intersect on the plane, then two of the points will be real and two of the points will be imaginary and complex conjugate – the two complex conjugate roots are the images of the circular points. If, on the other hand, all four points are imaginary, then there will be two pairs of complex conjugate roots. In this case, two transforms are possible to rectify the plane – in our experience, one transform yields a highly distorted plane, while the other yields the correctly rectified plane. Recall that the intersection points, Equation (20), are expressed in terms of the coefficients, α, β, and $l_\infty = \begin{pmatrix} l_1 & l_2 & l_3 \end{pmatrix}^T$, of the desired matrices P and A. These coefficients can now be directly computed as follows.

To make the solution algebraically more approachable, we begin by normalizing Equation (20) by $-il_3$ (this arbitrary scale factor does not affect the homogeneous coordinates):

$$H\vec{I} = \begin{pmatrix} \alpha l_3 - i\beta l_3 \\ l_3 \\ -l_2 - l_1\alpha + il_1\beta \end{pmatrix}. \quad (27)$$

Denote the estimated $H\vec{I}$ as $\vec{z} = \begin{pmatrix} x & y & 1 \end{pmatrix}^T$. Scaling $\vec{z}$ by the complex conjugate of y yields:

$$\vec{z} = \begin{pmatrix} xy^* \\ yy^* \\ 1y^* \end{pmatrix} = \begin{pmatrix} z_1 + iz_2 \\ z_3 \\ z_4 + iz_5 \end{pmatrix}. \quad (28)$$

Note that the first and third components are imaginary while the second is real, matching the form of $H\vec{I}$ in Equation (27). The desired parameters can, therefore, be determined simply as:

$$\alpha = z_1/z_3 \quad (29)$$
$$\beta = -z_2/z_3 \quad (30)$$
$$l_\infty = \begin{pmatrix} -z_5z_3/z_2 & z_5z_1/z_2 - z_4 & z_3 \end{pmatrix}^T \quad (31)$$

Figure 4: The face of each cigarette box is rectified using the known shape of the box face. Shown in the center and right panels are the rectified images, clearly showing an inconsistency in the cartoon character and text.

If the intersection point $H\vec{J}$ is used, instead of $H\vec{I}$, then the sign of β will be incorrect, and the resulting rectified plane will be reflected about the y-axis – this incorrect transform is typically visually obvious.

3 Metric Measurements

3.1 Polygons

Shown in Figure 1 is a car's rear license plate photographed at a grazing angle. The plate number, shown at full resolution in the upper-right panel of this figure, is largely illegible. Since 1957 the size of license plates have been standardized in the United States to 6×12 inches [6]. This known shape thus provides an ideal constraint for planar rectification from known polygons, Section 2.1. Shown in the lower-right panel of Figure 1 is the result of rectification (the image was also histogram equalized to improve overall readability). This result clearly reveals the plate number.

Shown in left-most panel of Figure 4 is a forgery of our creation – two boxes of Marlboro cigarettes were doctored to read "Marlboro kids" with an image of the cartoon character Tweety Bird. On both boxes, the "kids" text and the character were manually adjusted in Photoshop to give the appearance of correct perspective. Shown in the center and right panels of Figure 4 are the results of planar rectification based on the known shape of the rectangle on the front of the box ($1\ 11/16 \times 3\ 1/8$ inches, determined by measuring an actual box of cigarettes). Note that after rectification the text and character on the boxes are inconsistent with one another, clearly revealing them to be fakes. Interestingly, this incorrect perspective is not obvious in the original forgery.

3.2 Vanishing Lines

Shown in Figure 5 is an image of a crosswalk before (left) and after (right) rectification using vanishing lines, Section 2.2. The parallel lines along the crosswalk, and along the side and center of the

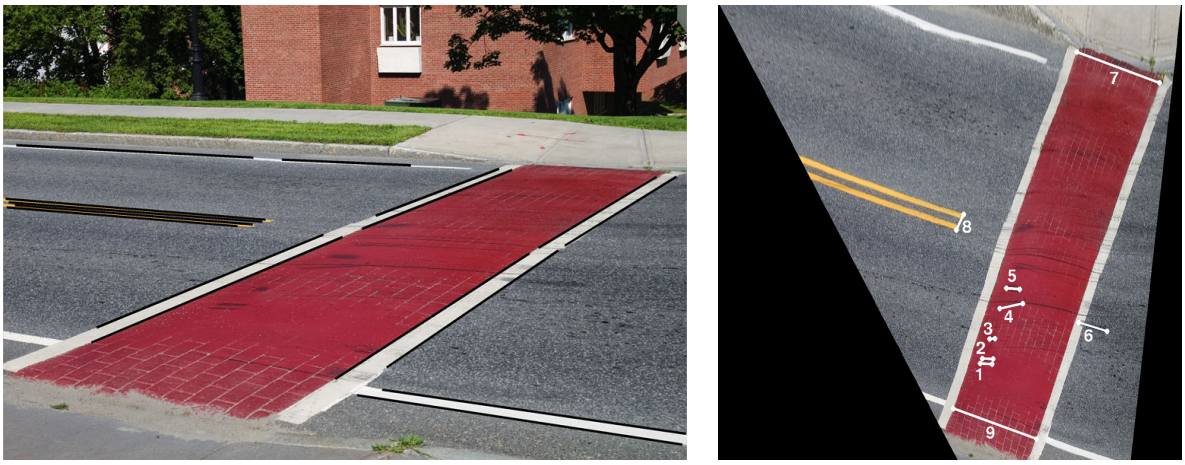

Figure 5: The road was rectified using vanishing lines and a known angle and length ratio. Shown on the right is the rectified image, from which measurements of skid marks can be made.

Figure 6: The ground plane was rectified using vanishing lines and a known angle and length ratio. Shown on the right is the rectified image, from which it can be seen that this image is most likely authentic.

street were used to determine a pair of vanishing points. While computing the vanishing point from just lines is straight-forward, a least-squares solution is employed to compute the vanishing point from more than two lines, Appendix F. The first constraint circle was determined by assuming that the angle between the white lines delineating the crosswalk and the side of the road was 90 degrees. The second constraint circle was determined from the known length ratio between twelve five-inch bricks (60 inches) and the thickness of one white line (4 inches). These measurements were also used to determine the final scaling matrix. The line segments labeled $1 - 9$ were physically measured on the road to be 8.0, 7.4, 3.5, 16.1, 9.0, 20.0, 59.5, 11.7 and 59.7 inches and measured in the rectified image to be 8.5, 8.0, 3.7, 16.7, 9.7, 20.7, 60.0, 11.5 and 60.0 inches. The average error between these measurements is 4.1%.

Shown in Figure 6 is an aerial image of an airplane taxiing on a runway and a seemingly unlikely advertisement on the neighboring grass. Also shown in this figure is the rectified version of this advertisement. The pair of parallel lines bounding the advertisement were used to determine vanishing

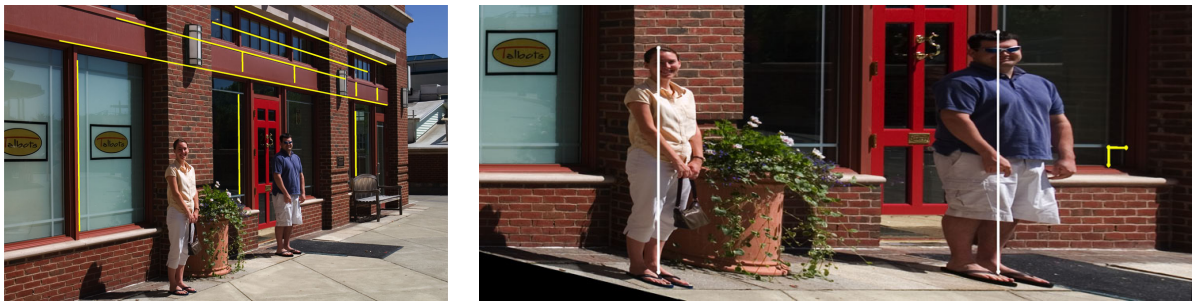

Figure 7: The building wall was rectified using vanishing lines and a known angle and length ratio. Shown on the right is the rectified image, from which measurements of the two people can be made.

points. The first constraint circle was determined by assuming that the angle between the top and right borders of the advertisement was 90 degrees. The second constraint circle was determined from the length ratio between the diagonals of the parallelogram connecting the letter "i" in "in" to the second "i" in "Virgin." This length ratio was assumed to be one. Note that while only an approximation, this constraint is reasonable for a parallelogram with one side much larger than the other, as in our case. The rectification is estimated up to an arbitrary scale factor. Note that the proportions of the text in the rectified image are consistent with a planar surface having undergone perspective projection. Were this a fake, it is likely that the text would have been placed with incorrect perspective distortion, as in Figure 4.

Shown in Figure 7 is an image of two people standing outside of a store. Also shown in this figure is a rectified version of this image. The lines along the building face were used to find vanishing points. The first constraint circle was determined by assuming the angle between the horizontal and vertical lines used for the vanishing point was 90 degrees. The second constraint circle was determined by an equal length ratio – the decorative border in the corners of the window is approximately a square. Since the two people are standing in a plane that is approximately parallel to the plane of the store front, their relative heights can be measured after rectification. Using the height of the person on the left as a reference (64.75 inches), the height of the person on the right was estimated to be 69.3 inches. This person's actual height is 68.75 inches, yielding an error of 0.55 inches or 0.8%.

3.3 Circles

Shown in Figure 8 is an image of a pair of parked cars before (top) and after (bottom) rectification using a pair of coplanar circles, Section 2.3. The rear wheel of the Chevy (left) and front wheel of the Honda (right) were used to rectify the image, and the known hubcap diameter of 16.5 inches was used to determine the final scaling matrix. The three line segments numbered 1, 2, and 3 denote the wheelbase of the Chevy, the distance between cars, and the wheelbase of Honda. These distances were measured in the physical world to be 106.6, 121.2 and 93.0 inches, and measured in the rectified image to be 107.1, 125.0 and 97.6 inches. The average error between these measurements is 2.7%.

Figure 8: The top image was rectified using known circles (wheels). Shown below is the rectified image from which the distance between the cars can be measured.

4 Discussion

This paper surveys three techniques for estimating the transformation, H, of a plane imaged under perspective projection. With this transformation, a planar surface can be rectified to be fronto-parallel. Each technique requires only a single image, and each exploits different geometric principles. In the first and most direct approach, the shape of a known polygon on the world plane is used to directly estimate H. In the second approach, vanishing lines and known angles or length ratios are used to estimate a projective and affine matrix, whose product yields H (up to a similarity transform). In the third, and final approach, a pair of coplanar circles are used to estimate the projective and affine matrix. In each approach, the estimation of H requires some user assistance in selecting lines, polygons, circles, etc.

We have shown that the removal of planar distortion can be helpful in certain forensic settings. In addition, under certain conditions real-world metric measurements can be made on the planar surface, providing another useful forensic tool.

5 Acknowledgments

This work was supported by a Guggenheim Fellowship, a gift from Adobe Systems, Inc., a gift from Microsoft, Inc., a grant (FA8750-06-C-0011) from the United States Air Force, and under a grant (2000-DT-CX-K001) from the U.S. Department of Homeland Security, Science and Technology Directorate (points of view in this document are those of the authors and do not necessarily represent the official position of the U.S. Department of Homeland Security or the Science and Technology Directorate).

Appendix A

Virtually all cameras introduce some amount of geometric distortion. These distortions are often described with a radially symmetric model [2, 8, 9], Figure 9. Given an ideal undistorted image $f(x, y)$, the distorted image is denoted as $f(\tilde{x}, \tilde{y})$, where the distorted spatial parameters are given by:

$$\tilde{x} = x(1 + \kappa r^2) + c_x \quad \text{and} \quad \tilde{y} = y(1 + \kappa r^2) + c_y, \tag{32}$$

where $r^2 = x^2 + y^2$, κ controls the amount of distortion, and (c_x, c_y) is the center of the distortion.

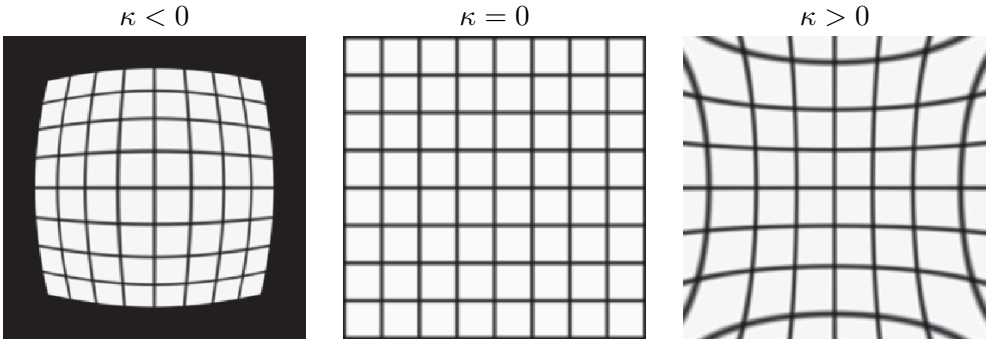

Figure 9: Radially symmetric lens distortion, Equation (32).

The lens distortion parameters can be determined by imaging a calibration target with known fiducial points. The deviation of these points from their original positions is used to estimate the amount of distortion (e.g., [9]). An alternative calibration technique relies on the presence of straight lines in the scene (e.g., [1, 7]). These lines, mapped to curves in the image due to the distortion, are automatically detected or specified by the user. The distortions are estimated by finding the model parameters that map these curved lines to straight lines. We adopt this latter approach since it is more realistic in a forensic setting.

A set of N arbitrarily spaced zero-meaned image coordinates along a curve are collected into the $N \times 2$ matrix D. The best-fit line, in a total least squares sense, to this data is given by a unit vector $\vec{u}$ that satisfies:

$$\min_{\vec{u}} \|D\vec{u}\|, \tag{33}$$

such that $\|\vec{u}\|^2 = 1$. The minimal eigenvalue eigenvector of $D^T D$ yields the optimal $\vec{u}$. The error between the best-fit line and the underlying image coordinates is used as a measure of "straightness":

$$\|D\vec{u}\|^2. \tag{34}$$

The lens distortion parameters are estimated using a brute-force search over the three model parameters κ, c_x, and c_y that minimizes the above measure of straightness averaged over all user specified curves.

An image is undistorted by solving Equation (32) for the original spatial coordinates x and y, and warping the distorted image onto this sampling lattice. Solving for the original spatial coordinates is done in polar coordinates where the solution takes on a more convenient form. In polar coordinates the undistorted image is denoted as $f(r, \theta)$, where:

$$r = \sqrt{x^2 + y^2} \quad \text{and} \quad \theta = \tan^{-1}(y/x). \tag{35}$$

Similarly, the distorted image $f(\tilde{x}, \tilde{y})$ in polar coordinates is $f(\tilde{r}, \tilde{\theta})$, where:

$$\tilde{r} = \sqrt{(\tilde{x} - c_x)^2 + (\tilde{y} - c_y)^2} \quad \text{and} \quad \tilde{\theta} = \tan^{-1}((\tilde{y} - c_y)/(\tilde{x} - c_x)). \tag{36}$$

Combining these parameters with Equation (32) and (35) yields:

$$\tilde{r} = r(1 + \kappa r^2) \quad \text{and} \quad \tilde{\theta} = \tan^{-1}(y/x). \tag{37}$$

Note that since the distortion model is radially symmetric, only the radial component is affected. The undistorted radial parameter r is determined by solving the resulting cubic equation in Equation (37), Appendix B. These polar parameters are then converted back to rectangular coordinates and the distortion is inverted by warping the image $f(\tilde{x}, \tilde{y})$ onto this new sampling lattice.

Appendix B

The roots of the cubic equation in r given by:

$$r^3 + a_2 r^2 + a_1 r + a_0 = 0 \tag{38}$$

are determined as follows. Let:

$$Q = (3a_1 - a_2^2)/9 \tag{39}$$
$$R = (9a_2 a_1 - 27a_0 - 2a_2^3)/54. \tag{40}$$

Let the polynomial discriminant be:

$$D = Q^3 + R^2. \tag{41}$$

If $D > 0$ then there is one real root and two complex conjugates. The real root is:

$$r_1 = -a_2/3 + (R + \sqrt{D})^{1/3} + (R - \sqrt{D})^{1/3}. \tag{42}$$

If $D \leq 0$ then all roots are real and given by:

$$r_1 = 2\sqrt{-Q}\cos(\theta/3) - a_2/3 \tag{43}$$
$$r_2 = 2\sqrt{-Q}\cos((\theta + 2\pi)/3) - a_2/3 \tag{44}$$
$$r_3 = 2\sqrt{-Q}\cos((\theta + 4\pi)/3) - a_2/3, \tag{45}$$

where:

$$\theta = \cos^{-1}(R/\sqrt{-Q^3}). \tag{46}$$

From these three roots, the root closest to the range $[0, 1]$ is chosen.

Appendix C

The intersection (α, β) of two circles with centers (c_x, c_y) and (d_x, d_y) and radii r_c and r_d is determined as follows:

$$\alpha = c_x + \frac{\gamma_2}{\gamma_1}(d_x - c_x) + \frac{\gamma_3}{\gamma_1}(d_y - c_y) \quad (47)$$

$$\beta = c_y + \frac{\gamma_2}{\gamma_1}(d_y - c_y) + \frac{\gamma_3}{\gamma_1}(d_x - c_x), \quad (48)$$

where,

$$\gamma_1 = \sqrt{(c_x - d_x)^2 + (c_y - d_y)^2} \quad (49)$$

$$\gamma_2 = (r_c^2 - r_d^2 + \gamma_1^2)/(2\gamma_1) \quad (50)$$

$$\gamma_3 = \sqrt{r_c^2 - \gamma_2^2}. \quad (51)$$

Appendix D

Any conic section, including ellipses, can be written as a quadratic equation in terms of image coordinates x and y:

$$c_1x^2 + 2c_2xy + c_3y^2 + 2c_4x + 2c_5y + c_6 \quad = \quad 0. \tag{52}$$

In homogeneous coordinates this equation can be expressed as:

$$\vec{x}^T C \vec{x} \quad = \quad 0\,, \tag{53}$$

with

$$C \quad = \quad \begin{pmatrix} c_1 & c_2 & c_4 \\ c_2 & c_3 & c_5 \\ c_4 & c_5 & c_6 \end{pmatrix}. \tag{54}$$

The conic coefficient matrix C for an ellipse can be estimated from $n \geq 5$ image coordinates by solving the following system of linear equations for the six coefficients of C:

$$\begin{pmatrix} x_1^2 & 2x_1y_1 & y_1^2 & 2x_1 & 2y_1 & 1 \\ x_2^2 & 2x_2y_2 & y_2^2 & 2x_2 & 2y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & 2x_ny_n & y_n^2 & 2x_n & 2y_n & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{pmatrix} \quad = \quad \vec{0}$$

$$M\vec{c} \quad = \quad \vec{0}. \tag{55}$$

The vector $\vec{c}$ that satisfies $\min_{\vec{c}} \|M\vec{c}\|$, such that $\|\vec{c}\|^2 = 1$, is the minimal eigenvalue eigenvector of M^TM.

Appendix E

Let $f(x) = a_2x^2 + a_1x + a_0$ and $g(x) = b_2x^2 + b_1x + b_0$. The resultant of these polynomials is defined to be:

$$\text{Res}(f, g) \quad = \quad (a_2b_1 - a_1b_2)(a_1b_0 - a_0b_1) - (a_2b_0 - a_0b_2)^2. \tag{56}$$

Theorem: The polynomials $f(\cdot)$ and $g(\cdot)$ have a common root if and only if:

$$\text{Res}(f, g) \quad = \quad 0. \tag{57}$$

This theorem and the following proof are a special case of a Theorem 5.7 in [4].

(**Proof:** $\Rightarrow$) If $f(x)$ and $g(x)$ have a common root, then there exists an x such that $f(x) = g(x) = 0$. As a result, arbitrary linear combinations of $f(x)$ and $g(x)$ must also be zero:

$$\begin{aligned} b_2f(x) - a_2g(x) &= 0 \\ b_2a_2x^2 + b_2a_1x + b_2a_0 - a_2b_2x^2 - a_2b_1x - a_2b_0 &= 0 \\ (b_2a_1 - a_2b_1)x + b_2a_0 - a_2b_0 &= 0 \\ x &= \frac{a_2b_0 - b_2a_0}{b_2a_1 - a_2b_1}\,, \end{aligned} \tag{58}$$

and

$$\begin{aligned} b_1f(x) - a_1g(x) &= 0 \\ b_1a_2x^2 + b_1a_1x + b_1a_0 - a_1b_2x^2 - a_1b_1x - a_1b_0 &= 0 \\ (b_1a_2 - a_1b_2)x^2 + b_1a_0 - a_1b_0 &= 0 \\ x^2 &= \frac{a_1b_0 - b_1a_0}{b_1a_2 - a_1b_2}\,. \end{aligned} \tag{59}$$

Combining Equations (58) and (59) yields:

$$\frac{a_1b_0 - b_1a_0}{b_1a_2 - a_1b_2} = \frac{(a_2b_0 - b_2a_0)^2}{(b_2a_1 - a_2b_1)^2} \tag{60}$$

$$\begin{aligned} (b_1a_0 - a_1b_0)(b_2a_1 - b_1a_2) &= (a_2b_0 - b_2a_0)^2 \\ (b_1a_0 - a_1b_0)(b_2a_1 - b_1a_2) - (a_2b_0 - b_2a_0)^2 &= 0 \\ \text{Res}(f, g) &= 0. \end{aligned} \tag{61}$$

(**Proof:** $\Leftarrow$) Assume that $f(x)$ and $g(x)$ satisfy Equation (61), and let x be given by Equation (58). Starting with Equation (61) and working in reverse, we can derive Equation (60) for x^2. It is then easy to verify that $f(x) = g(x) = 0$ using these values for x and x^2. The polynomials, $f(x)$ and $g(x)$, therefore, have a common root x.

Appendix F

Consider two lines parametrized as $a_1x + b_1y = c_1$ and $a_2x + b_2y = c_2$. The intersection of these lines (assuming that they are not parallel) can be determined by solving the following system of linear equations:

$$\begin{aligned} \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ M\vec{x} &= \vec{c} \\ \vec{x} &= M^{-1}\vec{c}. \end{aligned} \tag{62}$$

Given n lines, the intersection point that minimizes the least-squares distance to all of the lines can be determined by solving the following over-constrained system of linear equations:

$$\begin{aligned} \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \\ M\vec{x} &= \vec{c} \\ \vec{x} &= (M^TM)^{-1}M^T\vec{c}. \end{aligned} \tag{63}$$

To avoid bias in the estimator, each line is normalized so that $a_i^2 + b_i^2 = 1$ for $i \in [1, n]$.

References

[1] F. Devernay and O. Faugeras. Automatic calibration and removal of distortion from scenes of structured environments. In *SPIE Conference on Investigative and Trial Image Processing*, San Diego, CA, 1995.

[2] W. Faig. Calibration of close-range photogrammetric systems: Mathematical formulation. *Photogrammetric Eng. Remote Sensing*, 41(12):1479–1486, 1975.

[3] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.

[4] Nathan Jacobson. *Basic Algebra*. W.H. Freeman and Company, 1985.

[5] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. In *Computer Vision and Pattern Recognition*, pages 482–488, 1998.

[6] Special to The New York Times. Americas planning uniform auto tags. *New York Times*, page 21, Oct. 4, 1956.

[7] R. Swaminatha and S.K. Nayar. Non-metric calibration of wide-angle lenses and polycameras. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 413–419, Fort Collins, CO, 1999.

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

[9] J. Weng. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965–979, 1992.