

A Taxonomy and Analysis of Camera Calibration Methods for Traffic Monitoring Applications

Neeraj K. Kanhere and Stanley T. Birchfield, *Senior Member, IEEE*

Abstract—Many vision-based automatic traffic monitoring systems require a calibrated camera for computing speeds and length-based classifications of tracked vehicles. A number of techniques, both manual and automatic, have been proposed for performing such calibration, but no study has yet focused upon evaluating the relative strengths of these different alternatives. We present a taxonomy for roadside camera calibration that not only encompasses the existing methods (VWV, VWH, and VWL) but also includes several novel ones as well (VVH, VVL, VLH, VVD, VWD, and VHD). We also introduce an overconstrained approach that takes into account all the available measurements, resulting in reduced error as well as overcoming the inherent ambiguity in the single-vanishing-point solutions. This important but oft-neglected ambiguity has not received the attention that it deserves; we analyze it and propose several ways of overcoming it. Our analysis includes the relative tradeoffs between two-vanishing-point solutions, single-vanishing-point solutions, and solutions that require the distance to the road to be known. The various methods are compared using simulations and experiments with real images, showing that methods that use a known length generally outperform the others in terms of error, and that the overconstrained method reduces errors even further.

I. INTRODUCTION

AS the cost of cameras and processors continues to decrease, vision-based sensing is becoming an increasingly popular alternative to traditional sensors for collecting traffic data. A number of both research and commercial systems have shown the promise of gathering information such as volume, speed, classification, incidents, and turning movements from video. In addition to their cost, vision systems are especially attractive due to their non-intrusiveness and rich data gathering ability.

In order for a vision-based system to measure the speeds of vehicles, there must be a mapping from pixels in the image to coordinates in the world. Such a mapping can also be used to increase the accuracy of the estimation of other types of data such as volume, classification, or incidents. While some systems rely only upon pixel-based measurements to infer world lengths in certain directions [13], [21], a complete image-to-world mapping enables much more detailed analysis of the scene.

In recent years a number of authors have proposed various methods for calibrating a roadside camera. Trajković [17] describes an interactive approach to calibrating a Pan-Tilt-Zoom (PTZ) camera by assuming that the camera height is known. Bas and Crisman [1] use the known height and tilt angle of the camera, along with a single set of parallel lines (along the road edges) drawn by the user, while Lai [12] removes the assumption of known height and tilt angle by using an additional line of known length perpendicular to

the road edges. The technique of Fung et al. [5] uses the pavement markings and known lane width but requires the user to draw a rectangle formed by parallel lane markings in adjacent lanes. The problem of ill-conditioned vanishing points (i.e., parallel lines in the world appearing parallel in the image, so that the vanishing point is at infinity) has been addressed by He et al. [8] using known length and width of road lane markings. In an automatic approach, Schoepflin and Dailey [14] dynamically calibrate PTZ cameras using lane activity maps to find lane centers, along with gradient histograms to estimate the direction perpendicular to the direction of travel. Song et al. [15] use edge detection to find the lane markings in the static background image, from which the vanishing point is estimated by assuming that the camera height and lane width are known in advance. Note that these application-specific techniques are motivated by the fact that generic camera calibration techniques [18], [23] are difficult to apply to the roadside setting due to the lack of a calibration grid, the dominance of a single plane, and the variable focal length.

A natural question to ask is, What is the relationship between these methods? Early in our investigation we implemented an existing technique only to find that it was extremely sensitive to the input, so that even a slight amount of noise would cause the resulting estimates for the parameters to exhibit surprisingly large errors. This puzzling behavior led us to investigate the conditions under which such a technique would prove practically useful, as well as the conditions under which it cannot be trusted. We also sought to study the characteristics of the other techniques to see whether they would exhibit similar behavior, and we desired to consider related techniques that had not yet been proposed. In effect, our goal was to determine and analyze the space of roadside camera calibration methods.

In this paper we present the results of this investigation. Starting with a simplified pinhole camera model, we derive the relationship between 3D world coordinates and 2D image coordinates using the main parameters of the system. We then introduce a taxonomy for the different calibration methods that arise due to the differing image and world measurements available. We divide the methods into two categories, depending upon whether only a single vanishing point or two vanishing points are available. The taxonomy includes three existing techniques (VWV, VWH, and VWL) as well as six novel ones (VVH, VVL, VLH, VVD, VWD, and VHD). We show that under certain conditions the single-vanishing point techniques result in an ambiguous answer, and we introduce a heuristic to overcome the ambiguity. This ambiguity has not been previously analyzed nor given the attention that it deserves. In

addition, we propose an overconstrained approach that takes into account all the available measurements to refine the result of any of the individual methods as well as to overcome the ambiguity when additional measurements are available. Following this, we discuss how existing methods fit into the proposed taxonomy. Finally, we present simulations and experiments to quantitatively compare the different methods under a variety of conditions, showing the improvement that results from the overconstrained approach.

II. THE CAMERA MODEL

Consider a pinhole camera viewing a straight, flat road. Assume that the camera has zero roll angle, square pixels (i.e., unity aspect ratio), zero skew, and a principal point at the image center, as in [14], [15]. With these assumptions, exactly three parameters are needed to map Euclidean points on the road plane to image points on the image plane: the camera focal length f , the height h of the center of projection above the road plane, and the tilt angle ϕ . In addition, the pan angle θ is needed to align the road coordinates with the direction of traffic flow. Note that a non-zero roll angle (also known as the “swing angle”) can be compensated by a simple image rotation, without affecting the findings of this paper.

Let us define two coordinate systems, as shown in Figure 1. The camera coordinate system is centered on the image plane at the principal point, with x_c and y_c axes that are aligned with the rows and columns of the image, respectively. (Using the standard image convention, the x_c axis points toward the right of the image, while the y_c axis points downward.) The world coordinate system is centered at the point on the road plane directly beneath the center of projection. The x axis is parallel to the x_c axis, while the y axis is perpendicular to the x axis so that, in a right-handed coordinate system, the z axis points upward from the road. We adopt the convention that $\phi = 0$ when the camera is aimed at the horizon (i.e., the optical axis z_c is parallel to the y axis), so that the two coordinate frames are related by a rotation of $\phi + \frac{\pi}{2}$ about the x axis. The pan angle θ is defined as the angle between the yz plane and the direction of travel.

The projection of a point (x, y, z) in the world coordinate frame to a point (u, v) in the image can be expressed as

$$\mathbf{p} = P\mathbf{x} = KRT\mathbf{x}, \quad (1)$$

where the homogeneous coordinates of the world point and its projection are $\mathbf{x} = [x \ y \ z \ 1]^T$ and $\mathbf{p} = [\alpha u \ \alpha v \ \alpha]^T$, $\alpha \neq 0$, respectively, and (u, v) is a scaled version of (x_c, y_c) according to the focal length f of the camera. The matrix

$$K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

captures the internal camera parameters,

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sin \phi & -\cos \phi \\ 0 & \cos \phi & -\sin \phi \end{bmatrix} \quad (3)$$

corresponds to a rotation of $\phi + \frac{\pi}{2}$ about the x axis, and

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -h \end{bmatrix} \quad (4)$$

captures the vertical translation of the camera. Note that f is in pixels, h in meters, $0 < \phi < \frac{\pi}{2}$, and $0 \leq \theta \leq \pi$. Expanding (1) yields the imaging expression for our model:

$$\begin{bmatrix} \alpha u \\ \alpha v \\ \alpha \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & -f \sin \phi & -f \cos \phi & fh \cos \phi \\ 0 & \cos \phi & -\sin \phi & h \sin \phi \end{bmatrix}}_{P=KRT} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \quad (5)$$

where $P = KRT$ is the 3×4 projection matrix. Converting to inhomogeneous coordinates, we see that a point $(x, y, 0)$ on the road plane projects onto the image at

$$u = \frac{\alpha u}{\alpha} = \frac{fx \sec \phi}{y + h \tan \phi} \quad (6)$$

$$v = \frac{\alpha v}{\alpha} = \frac{fh - fy \tan \phi}{y + h \tan \phi}. \quad (7)$$

These are the fundamental equations used to describe the mapping from the road plane to the image plane. In the following sections we describe a number of different scenarios in which the calibration parameters f , h , ϕ , and θ can be estimated.

A. Two vanishing points

Using the convention described above, the point at infinity associated with the direction of traffic flow along the length of the road is represented in 3D homogeneous coordinates as $\mathbf{x}_0 = [-\tan \theta \ 1 \ 0 \ 0]^T$. Similarly, the point at infinity associated with the direction in the road plane perpendicular to traffic flow is $\mathbf{x}_1 = [1 \ \tan \theta \ 0 \ 0]^T$. Projecting onto the image plane yields the vanishing points (u_0, v_0) and (u_1, v_1) : $[\alpha u_i \ \alpha v_i \ \alpha]^T = P\mathbf{x}_i$, where $i \in \{0, 1\}$. (Note that the arbitrary scaling factor $\alpha \neq 0$ need not be related in the various equations in which it appears.) By expanding this expression, we obtain

$$u_0 = \frac{\alpha u_0}{\alpha} = \frac{-f \tan \theta}{\cos \phi} \quad (8)$$

$$v_0 = \frac{\alpha v_0}{\alpha} = \frac{\alpha v_1}{\alpha} = -f \tan \phi \quad (9)$$

$$u_1 = \frac{\alpha u_1}{\alpha} = \frac{f}{\cos \phi \tan \theta}, \quad (10)$$

and $v_0 = v_1$ arises from the zero roll angle assumption. It is straightforward to solve these three equations for the unknowns:

$$f = \sqrt{-(v_0^2 + u_0 u_1)} \quad (11)$$

$$\phi = \tan^{-1} \left(\frac{-v_0}{f} \right) \quad (12)$$

$$\theta = \tan^{-1} \left(\frac{-u_0 \cos \phi}{f} \right). \quad (13)$$

Note several facts:

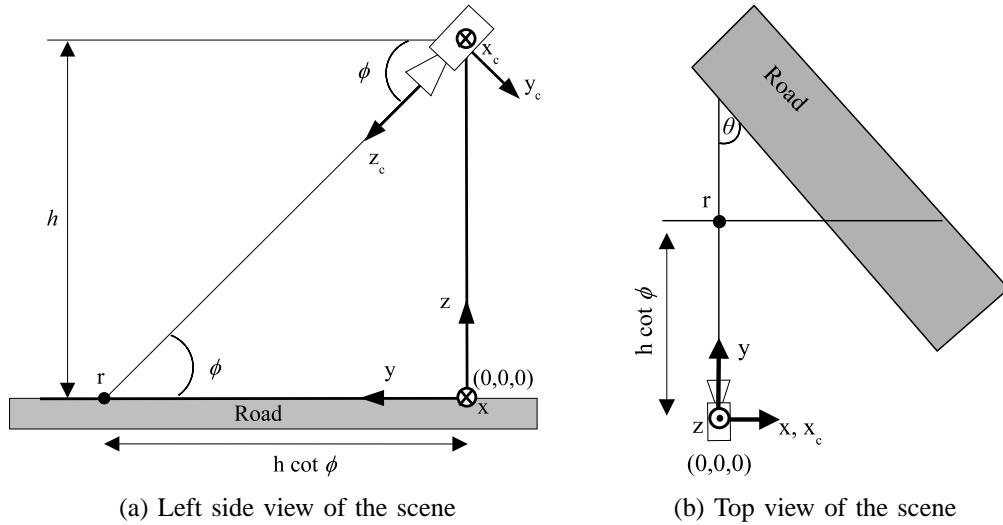


Fig. 1. The coordinate systems used. The camera is placed at a height of h feet above the road with tilt angle ϕ and pan angle θ . The world coordinate system is defined by the x , y , and z axes, while x_c and y_c define the camera coordinate system. Both systems are right-handed, so that x and x_c (which are parallel to each other) are into the page in the left side view, as indicated by the symbol \otimes , while z is coming out of the page in the top view, shown by \odot . The optical axis of the camera z_c intersects the road plane along the y axis at $r = (0, h \cot \phi, 0)$.

- The height h does not appear in the equations, because vanishing points are independent of the camera position. In fact, dividing (8) by (9) reveals that the vanishing point coordinates bear a simple relation to the angles:

$$\frac{u_0}{v_0} = \frac{\tan \theta}{\sin \phi} \quad (14)$$

$$\frac{u_1}{v_1} = \frac{-1}{\sin \phi \tan \theta}. \quad (15)$$

- $v_0 \leq 0$. In other words, because the camera is pointed downward ($\phi \geq 0$), the horizon line is in the top half of the image. Similarly, u_0 is negative when θ is an acute angle and positive when θ is obtuse:

$$u_0 \leq 0 \quad \text{if} \quad 0 < \theta < \frac{\pi}{2} \quad (16)$$

$$u_0 \geq 0 \quad \text{if} \quad \frac{\pi}{2} < \theta \leq \pi, \quad (17)$$

where $u_0 \rightarrow \pm\infty$ as $\theta \rightarrow \frac{\pi}{2}$.

- The focal length f will always be real, because the term under the radical in (11) is guaranteed to be positive. This can be seen by substituting (8)-(10) into the expression $-u_0 u_1 > v_0^2$, which leads to $\sin^2 \phi < 1$, which is always true.

We will now consider three different scenarios for calibration using both vanishing points.

1) *Two vanishing points and known camera height (VVH)*: Often the height of the camera is already known, e.g., it was measured at installation time. When a camera pans, tilts, and zooms, its height above the ground does not change. The simplest scenario therefore is to measure the two vanishing points (u_0, v_0) and (u_1, v_1) and then compute the parameters f , θ , and ϕ using (11)-(13). Since h is already known, the calibration is complete.

2) *Two vanishing points and known width (VWV)*: If, however, the height h is unknown, it can be computed using a single known distance in the road perpendicular to the

direction of traffic flow. To develop the solution, note from (6) that for any two points in the road plane having the same coordinates along the y axis, their distance Δx in the world is related to the difference Δu in the image coordinates of their projections by

$$\Delta u = \frac{f \Delta x}{y \cos \phi + h \sin \phi}. \quad (18)$$

Solving for h yields

$$h = \frac{1}{\sin \phi} \left(f \frac{\Delta x}{\Delta u} - y \cos \phi \right). \quad (19)$$

Now let us consider the special case of a width measured on the u axis in the image. By substituting $v = 0$ in (7) and solving for y , we obtain

$$y|_{v=0} = h \cot \phi. \quad (20)$$

From this result, we see that the optical axis intersects the road plane at the point $\mathbf{r} = (0, h \cot \phi, 0)$, shown in Figure 1. Substituting into (19) yields an expression for the camera height

$$h = f \sin \phi \left. \frac{\Delta x}{\Delta u} \right|_{v=0} \quad (21)$$

given a known distance Δx of a line segment on the road at $y = h \cot \phi$ and the image distance Δu of the corresponding projection onto the image along the row $v = 0$.

As shown in Figure 2, a width w perpendicular to the traffic flow produces a distance $\Delta x = w \sec \theta$ along any line parallel to the x axis. Typical values that can be measured are the width of a vehicle or the lane width. If either of these is known, and if the corresponding image measurement Δu can be made, then the height h can be computed using

$$h = \frac{f w \sin \phi}{\delta \cos \theta}, \quad (22)$$

where $\delta = \Delta u|_{v=0}$ is the horizontal length of the projected line segment in the image along the $v = 0$ line. As before, the parameters f , θ , and ϕ are computed using (11)-(13).

Keep in mind that the two widths are used in slightly different ways. The δ corresponding to the road width can be measured from the intersection of the lane lines with the $v = 0$ axis in the image. However, the δ corresponding to the vehicle width is obtained by projecting the vehicle width onto the $v = 0$ axis using the first vanishing point (u_0, v_0) .

3) *Two vanishing points and known length (VVL)*: Suppose that instead of knowing the width of a line segment perpendicular to the road, we know the length of a line segment parallel to the road. To incorporate length information, we will use the v coordinates of the projection. To do this, we first rearrange (6) and (7) to get

$$x = \frac{h u \sec \phi}{v + f \tan \phi} \quad (23)$$

$$y = \frac{h(f - v \tan \phi)}{v + f \tan \phi}. \quad (24)$$

Note that y is independent of u , because the roll angle is zero. Now suppose there is a line segment on the road plane parallel to the direction of traffic flow with length ℓ , as shown in Figure 2. Typically this would be either the distance between pavement markings or the length of a vehicle. The front point lies at $(\cdot, y_f, 0)$ and projects onto the image at (\cdot, v_f) , where the dot indicates that the value is irrelevant to the computation. The back point lies at $(\cdot, y_b, 0)$, where $y_b = y_f + \ell \cos \theta$, and projects onto the image at (\cdot, v_b) .

Substituting into (24) yields

$$y_f = \frac{h(f - v_f \tan \phi)}{v_f + f \tan \phi} \quad (25)$$

$$y_f = \frac{h(f - v_b \tan \phi)}{v_b + f \tan \phi} - \ell \cos \theta. \quad (26)$$

Equating these two expressions, substituting $\tan \phi = -v_0/f$ from (9), and solving for h yields an expression for the height:

$$h = \frac{f \kappa \ell \cos \theta}{f^2 + v_0^2}, \quad (27)$$

where $\kappa = (v_f - v_0)(v_b - v_0)/(v_f - v_b)$. Again, (11)-(13) are used to compute f , θ , and ϕ .

B. One vanishing point

While estimating the vanishing point (u_0, v_0) in the direction of traffic flow is generally straightforward, it is often quite difficult to obtain an accurate estimate of the vanishing point (u_1, v_1) in the perpendicular direction. Worse, when $u_1 = \infty$ (which occurs when $\theta = 0$), the three approaches just described will not work because (11) is rendered useless. As a result, we seek a calibration solution that does not require the latter vanishing point. From (12) and (13), it is straightforward

to obtain the following:

$$\sin \phi = \frac{-v_0}{\sqrt{f^2 + v_0^2}} \quad (28)$$

$$\cos \phi = \frac{f}{\sqrt{f^2 + v_0^2}} \quad (29)$$

$$\sin^2 \theta = \frac{u_0^2}{f^2 + u_0^2 + v_0^2} \quad (30)$$

$$\cos^2 \theta = \frac{f^2 + v_0^2}{f^2 + u_0^2 + v_0^2}. \quad (31)$$

We will now use these equations to derive three additional scenarios.

1) *One vanishing point, known width, and camera height (VWH)*: Squaring both sides of (22), substituting (28) and (31), and rearranging yields a fourth-order equation in f :

$$(1 - k_W^2) f^4 + [2v_0^2 - k_W^2(u_0^2 + v_0^2)] f^2 + v_0^4 = 0, \quad (32)$$

where $k_W = wv_0/h\delta$. Thus, with a single vanishing point, the height of the camera, and a known width, we can solve this equation to get the square of the focal length. When the equation yields exactly one positive root, f can be determined uniquely.

Otherwise, when the equation yields two positive values for f^2 , there is an ambiguity in determining f . This ambiguity will be analyzed in more detail in the next section. Note that in the special case that $u_0 = 0$, the ambiguity disappears, because (32) reduces to

$$[(1 - k_W^2) f^2 + v_0^2] (f^2 + v_0^2) = 0. \quad (33)$$

Since the right factor can never be zero for a real camera, the left factor must be zero. Rearranging therefore yields a unique solution for f in this case:

$$f = \sqrt{\frac{-v_0^2}{1 - k_W^2}}. \quad (34)$$

Note that this expression could also be derived by squaring both sides of (22) and substituting (28), noting that $u_0 = 0$ implies $\cos \theta = 1$ from (31).

Once f is found, the parameters ϕ and θ are computed using (12)-(13).

2) *One vanishing point, known length and camera height (VLH)*: Now suppose that instead of a known width, we have a known length. This information can be exploited in a similar manner by squaring both sides of (27), substituting (31), and rearranging to obtain another fourth-order equation in f :

$$f^4 + [u_0^2 + 2v_0^2 - k_L^2] f^2 + (u_0^2 + v_0^2)v_0^2 = 0, \quad (35)$$

where $k_L = \kappa \ell / h$. This equation always has two positive roots, leading to two possible values for the focal length, as we shall see. Once the correct f is selected, (12)-(13) can be used to find ϕ and θ .

3) *One vanishing point, known width and length (VWL)*: Now suppose that both a width and length are known, but the height of the camera is unknown. By equating (22) and (27), we derive yet another fourth order equation in f :

$$f^4 + [2(u_0^2 + v_0^2) - k_V^2] f^2 + (u_0^2 + v_0^2)^2 - k_V^2 v_0^2 = 0, \quad (36)$$

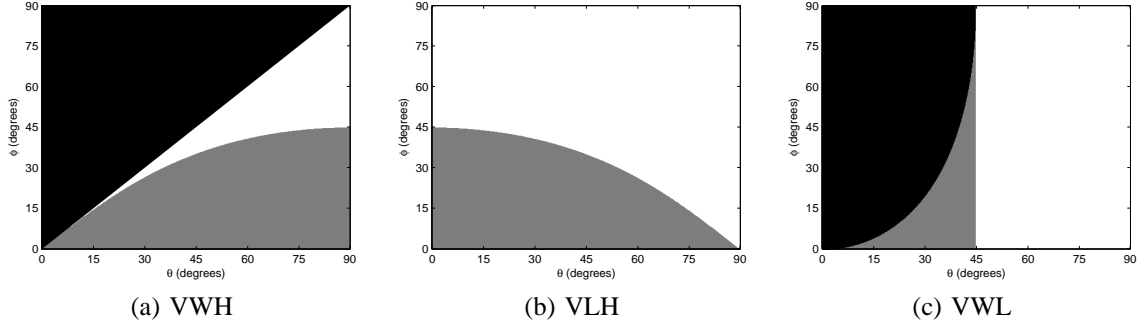


Fig. 3. Using a single vanishing point requires solving a quadratic equation for $m = f^2$. Depending on the pan (θ) and tilt (ϕ) angles, the equation can sometimes be solved uniquely for f because exactly one root for m is positive (black region). Otherwise, both solutions are positive and an ambiguity exists, in which case either m_+ yields the correct solution (gray region) or m_- yields the correct solution (white region). In the black region for VWH, the correct solution is m_- , while for VWL it is m_+ . All three plots form a mirror reflection about $\theta = 90^\circ$ for obtuse pan angles.

by noticing that when the solution in either (39) or (40) is differentiated with respect to one of the coefficients, the term $\sqrt{b^2 - 4ac}$ ends up in the denominator, thus leading to theoretically infinite sensitivity along the curve.

To summarize, if $\text{sgn}(m_+) \neq \text{sgn}(m_-)$, then we simply select $f = \sqrt{m_+}$ if $m_+ > 0$ or $f = \sqrt{m_-}$ if $m_- > 0$. Otherwise, if both m_+ and m_- are positive, then there is an ambiguity, in which case it is impossible without additional information to say whether m_+ or m_- is the correct solution. We suggest overcoming this ambiguity by either using the distance to the road described in the next section, the over-constrained method introduced in Section V, or the heuristic introduced in Section VII-B.

IV. USING DISTANCE FROM THE ROAD

Often the perpendicular distance d from the camera to the edge of the road, in the road plane, is known. From Figure 2 the following relationships are evident:

$$y_4 = d \csc \theta \quad (46)$$

$$y_2 = h \cot \phi \quad (47)$$

$$\tan \theta = \frac{x_2}{y_4 - y_2}. \quad (48)$$

While this parameter is not part of the camera model itself, as seen in (5), its availability leads to four additional scenarios for calibration.

A. Two vanishing points, known distance to the road (VVD)

Using (46)-(48) and substituting $u = u_2$, $x = x_2$ and $y = y_2$ in (6) yields an expression for the height in terms of d :

$$h = \frac{f d \tan \phi}{u_2 \sec \phi \cos \theta + f \sin \theta}. \quad (49)$$

Using two vanishing points, we already know f , ϕ , and θ from (11)-(13). Therefore, this expression completes the calibration if d is known. Note that this approach, as with the other techniques that use two vanishing points, cannot be used when $\theta = 0$ because $u_1 = \infty$.

Additionally, (49) can be used to overcome the ambiguity of the single-vanishing point techniques already discussed (VWH, VLH, and VWL), if d is known. Of the two solutions, the one that satisfies this equation is the correct solution.

B. One vanishing point, known width and distance (VWD)

Equating (22) with (49) and rearranging yields:

$$f \tan \theta \cos \phi = \frac{\delta d}{w} - u_2. \quad (50)$$

From (8) and (29) we have

$$\tan \theta \cos \phi = \frac{-u_0 f}{f^2 + v_0^2}. \quad (51)$$

Substituting into (50) and solving for f yields

$$f = \sqrt{\frac{k_D v_0^2}{1 - k_D}}, \quad (52)$$

where $k_D = (w u_2 - \delta d) / w u_0$. Thus, given a single vanishing point, a known width in the world, and the distance d , this equation allows us to compute f , from which ϕ and θ are determined as before. The height h can then be computed using either (22) or (49).

Note that this approach also does not work when $\theta = 0$, because (22) and (49) lead to a cancellation of f :

$$h = f \sin \phi \frac{w}{\delta} = f \sin \phi \frac{d}{u_2} \Rightarrow w u_2 = d \delta.$$

As a result, $k_D = 0/0$ is indeterminate, rendering (52) useless. In other words, when $\theta = 0$ the information provided by the width w and the distance d are the same, thus making h or ℓ necessary to solve for f .

C. One vanishing point, known height and distance (VHD)

Rearranging (49) and substituting from (9) yields

$$f \sin \theta + u_2 \sec \phi \cos \theta = \frac{-d v_0}{h},$$

which, after substituting $\sec \phi = \frac{-u_0}{f \tan \theta}$ from (8), becomes

$$f^2 \sin^2 \theta - u_0 u_2 \cos^2 \theta = \frac{-d v_0 f}{h} \sin \theta. \quad (53)$$

Squaring both sides and using (30) and (31) leads to a fourth-order equation in f :

$$\left[\frac{(u_2 - u_0)^2}{v_0^2} - \frac{d^2}{h^2} \right] f^4 + \left[2(u_2 - u_0)u_2 - \frac{d^2(u_0^2 + v_0^2)}{h^2} \right] f^2 + u_2^2 v_0^2 = 0. \quad (54)$$

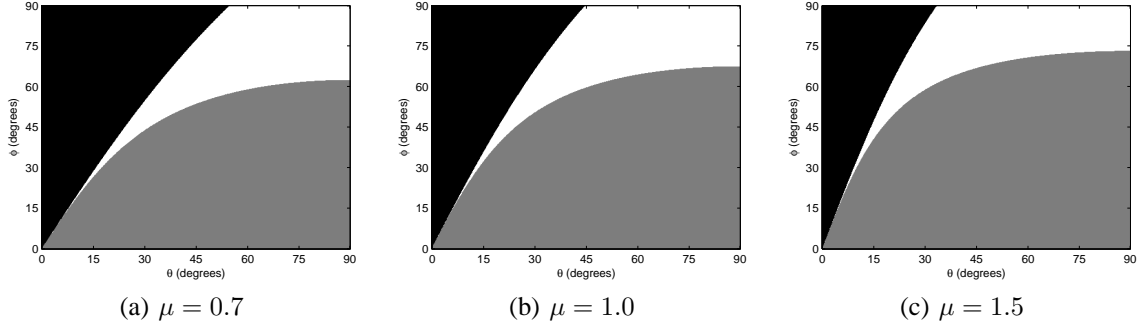


Fig. 4. The solutions for VHD follow a similar trend to those of VWH in Figure 3. In all cases $d = 30$ feet, and $h \in \{21, 30, 45\}$ feet, leading to $\mu \in \{0.7, 1.0, 1.5\}$.

Thus with a single vanishing point, known height and distance, this equation can be solved for f^2 . However, like before, this approach leads to an ambiguity if both roots are positive.

Unlike the previous cases of ambiguity (VWH, VLH, VWL), the ambiguity of VHD also depends upon the ratio $\mu = h/d$. Figure 4 shows the behavior of the solutions of (54) as a function of ϕ and θ for three different values of μ . Note that the location and shape of the regions remain largely unaffected by even substantial changes in μ . To determine the unambiguous black region, note that as in the previous cases involving a single vanishing point, VHD yields a single real solution when $ac < 0$, or

$$\frac{d^2}{h^2} > \frac{(u_2 - u_0)^2}{v_0^2}. \quad (55)$$

Solving (49) for u_2 and substituting, along with (8) and (9), yields an expression for the unambiguous region:

$$\cos \theta - \cos \phi > \mu \sin \theta \sin \phi. \quad (56)$$

Similarly, the boundary between the gray and white regions in the figure is given by $b^2 = 4ac$, which leads to

$$4(\cos \phi + \mu \sin \phi \sin \theta) \psi_{\phi, \theta} - \psi_{\phi, \theta}^2 = 4 \cos^2 \theta, \quad (57)$$

where

$$\psi_{\phi, \theta} = \frac{\sin^2 \theta + \sin^2 \phi \cos^2 \theta}{\sin \phi \cos \phi (\sin \phi - \mu \sin \theta \cos \phi)}. \quad (58)$$

D. One vanishing point, known length and distance (VLD)

Equating (27) and (49) yields, after some manipulation, a sixth-order equation in f :

$$\begin{aligned} f^6 + [2u_0^2 + 3v_0^2 - \frac{\gamma^2}{v_0^2}(u_0 - u_2)^2]f^4 \\ + [(u_0^2 + v_0^2)^2 + 2v_0^2(u_0^2 + v_0^2) + 2\gamma^2 u_2(u_0 - u_2)]f^2 \\ + [v_0^2(u_0^2 + v_0^2) - \gamma^2 v_0^2 u_2^2] = 0, \end{aligned} \quad (59)$$

where $\gamma = \kappa \ell / d$. This is the worst ambiguity of all, yielding up to three solutions for f . As a result, in our experiments we exclude this technique.

V. OVERCONSTRAINED METHOD

Sometimes more measurements are available than are needed by any method described so far. For example, in addition to the required quantities of VWL, we might also know the height and the distance to the road, and, in fact, we might even have multiple known lengths or widths in the image. To take advantage of such additional information, we propose an overconstrained (OC) method that minimizes an energy functional composed of the known distance to the road \bar{d} , known height \bar{h} , and known measurements (e.g., lengths or widths) in the road plane $\bar{\omega}_i$. That is, we find

$$\arg \min_{f, h, \phi} \lambda_d |d - \bar{d}| + \lambda_h |h - \bar{h}| + \lambda_\omega \sum_{i=1}^n |\omega_i - \bar{\omega}_i|, \quad (60)$$

where n is the number of known measurements in the road plane, and λ_d , λ_h , and λ_ω are scaling factors that weigh the importance of the different terms. (We set $\lambda_d = \lambda_h = \lambda_\omega = 1$.) Since the functional can include any number of measurements in the road plane, it is possible to achieve even greater redundancy by using image distances of multiple vehicles, along with a priori distributions of known lengths and widths of vehicles, with the lengths perhaps categorized by the type of vehicle. At any rate, the functional is minimized by applying Levenberg-Marquardt non-linear optimization starting from the parameters given by VWL (or any other method). In the case of ambiguity, all sets of parameters are used as starting points. In this way, this approach can not only improve the quality of the final solution, but it can also overcome the ambiguity.

Note that the optimization finds the parameter values that minimize the distance measurement error, perhaps at the expense of increasing the error in the estimated parameters themselves. This is a desirable outcome, however, since the primary goal of camera calibration is to measure distances in the road plane. In contrast, generic camera calibration approaches generally minimize image reprojection error using some sort of bundle adjustment.

VI. TAXONOMY APPLIED TO EXISTING METHODS

Over the years various traffic researchers have proposed methods for camera calibration, both manual and automatic. Table I illustrates how these methods fit into our taxonomy.

Note that the table does not include simplified calibration methods that do not yield the camera parameters [2], [3].

A number of authors have proposed solutions that fall within the category of VVW, which has been by far the most popular method to date. Lai and Yung [12] and Schoepflin and Dailey [14] both describe automatic methods for finding the lane boundaries and a line perpendicular to these lines (i.e., perpendicular to the direction of traffic flow). The lane boundaries are used to compute the first vanishing point, while the assumption of zero roll angle is used to estimate the second vanishing point from the perpendicular line. In our previous work we used this same formulation for both manual [10] and automatic [11] calibration. Once calibration has been performed, Schoepflin and Dailey [14] mention that the calculated h and d can be used to recalibrate after PTZ movement using a single vanishing point (VHD), though no results are shown.

The zero roll angle assumption is not invoked by Fung et al. [5], who require the user to click on four points corresponding to a rectangle in the road and then compute the two vanishing points from the two sets of parallel lines. Similarly, Zhaoxue and Pengfei [24] propose a particularly interesting method for overcoming a non-zero roll angle. They obtain the vanishing line by geometric construction using three parallel lines of equal intervals (e.g., the boundaries of two adjacent lanes with equal widths). Both of these approaches can be classified as VVW.

To avoid the difficulty of estimating the second vanishing point, several researchers have proposed single-vanishing point solutions. Gupte et al. [7] describe an interactive user interface in which the camera parameters are solved using an iterative nonlinear minimization technique assuming known quantities similar to those required by VWL. He and Yung [8] analyze the problem of ill-conditioned vanishing points (i.e., parallel lines in the world appearing parallel in the image, which is often the case for the second vanishing point). Their approach assumes that two sets of parallel world lines are known in the image, one pointing along the direction of traffic flow and the other perpendicular. If either set is nearly parallel in the image then the other set is used to find the other vanishing point in a manner similar to VWL, but if neither set is nearly parallel in the image then their method reduces to the VVW approach of Fung et al. [5]. We borrow much of our notation from Song et al. [15], who estimate the vanishing point automatically using edge detection to find the lane markings in the static background image, from which the calibration parameters are determined using VWH. Their paper is especially relevant because it not only introduces the second-order equation for m but also explicitly mentions that they use d to overcome the ambiguity in the solution.

Some previous methods require prior knowledge of camera parameters. For example, in the early work of Bas and Crisman [1], both the height and the tilt angle of the camera are assumed known beforehand. The vanishing point is computed from two parallel lines manually drawn in the image along the lane boundaries. From these quantities, the other parameters can easily be estimated using (11)-(13). Similarly, the method of Wu et al. [20] requires the focal length to be known

beforehand. Using two parallel lines found automatically in the image by edge detection, along with a known width, the other parameters can be found using (11)-(13) and (22). These methods, neither of which allows for non-zero roll angle, could be included in the table by expanding our taxonomy to include VHF and VWF, respectively.

A recent solution presented by Zhang et al. [22] relies on estimating vanishing points in three orthogonal directions, using pedestrians to obtain the vertical vanishing point. This work is reminiscent of the well-cited work of Wang and Tsai [19], which also use three vanishing points albeit in the plane. By solving for the complete camera calibration matrix, the approach of [22] solves for the principal point and roll angle, in what could be termed VVVH, since the height of the camera is known.

VII. EXPERIMENTAL RESULTS

The various calibration methods of our taxonomy were evaluated using both simulations and real experiments.

A. Simulations

A simulated camera satisfying the assumptions described in the beginning of Section II was placed near a simulated straight, flat stretch of road. The camera was placed at three different heights (30, 40 and 50 feet) above the road, and at three different distances from the edge of the road (40, 20, and -20 feet), where the third distance indicates that the camera was in the middle of the road. The tilt and pan angles of the camera were changed from 10° to 60° and 0° to 50° , respectively, in 10° increments. Three focal lengths were selected (250, 300, and 500 pixels), leading to an exploration space consisting of 972 configurations ($3 \text{ focal lengths} \times 3 \text{ heights} \times 3 \text{ distances} \times 6 \text{ tilt angles} \times 6 \text{ pan angles}$).

For each of these configurations we conducted 2000 simulations with noisy measurements to analyze the effect of error on the estimated camera parameters. Both image and world measurements were perturbed by adding random Gaussian noise. More precisely, we set $\tilde{q} = q + \xi$, where q and \tilde{q} represent the original and perturbed measurements, respectively, and $\xi \sim \mathcal{N}(0, \sigma^2)$ is a zero-mean normal random variable with variance σ^2 .

Prior to the simulation, a set of line segments in the road plane was randomly generated within the field of view, and a ground truth simulated camera was used to yield the corresponding image coordinates for each line segment. These values were used to test the accuracy of the parameter estimation as follows. After a given method was used to estimate the camera parameters from the noisy measurements, the image coordinates of the line segments were backprojected onto the road plane using the estimated parameters. The estimated distance along each line segment was compared with the corresponding ground truth distance, and the mean absolute error in these estimated distances was used to quantify the accuracy of the method.

Note that our evaluation is directly based upon measurements in the road plane. This decision is due to the specific application being considered, in which the purpose of the

| Previous work | Method | Known quantities | Non-zero roll? |
|----------------------------|--------|--|----------------|
| Lai and Yung [12] | VVW | Two parallel lines, one perpendicular line, and lane width | No |
| Schoepflin and Dailey [14] | VVW | Multiple parallel lines, one perpendicular line, and lane width | No |
| Kanhere et al. [10] | VVW | Two parallel lines, one perpendicular line, and lane width | No |
| Kanhere et al. [11] | VVW | Two parallel lines, one perpendicular line, and vehicle width | No |
| Fung et al. [5] | VVW | Rectangular pattern of known width | Yes |
| Zhaoxue and Pengfei [24] | VVW | Three parallel lines with identical known widths, and perpendicular line | Yes |
| Gupte et al. [7] | VWL | Two parallel lines, lane width, distance between lane markings | Yes |
| He and Yung [8] | VWL | Rectangular pattern of known width and length | Yes |
| Song et al. [15] | VWH | Lane width and height of the camera | No |

TABLE I

PREVIOUS WORK CATEGORIZED USING THE PROPOSED TAXONOMY OF CALIBRATION METHODS. THE LAST COLUMN INDICATES WHETHER EACH METHOD ALLOWS FOR A NON-ZERO ROLL ANGLE.

camera is solely to obtain speeds and other information with respect to the road. Our approach is in contrast to evaluations of generic camera calibration methods based upon image measurements, such as [6], [25], where the primary focus is upon recovering internal camera parameters of fixed-focal length cameras. Other researchers [4], [9], [16] compare generic calibration techniques using sophisticated equipment such as laser range finders or structured light systems. Such equipment would be difficult to apply in the application considered due to the dynamic nature of the scene and the outdoor illumination levels.

Figures 5 and 6 show simulation results for four different angles for the values $f = 500$ pixels, $d = 20$ feet, and $h = 50$ feet. Notice from these figures that, in general, the sensitivity to measurement errors decreases significantly as the tilt angle ϕ increases to a moderate value. Note that the methods based on length (specifically VVL, VWL, and VLH) outperformed the other methods, and that the two-vanishing-point methods were generally more accurate than those using a single vanishing point. The other simulations, which are omitted due to space limitations, exhibit similar qualitative trends as those shown in these figures. To give the reader an idea of this stability, Figure 7 shows the results when the focal length and ratio $\mu = h/d$ is changed. The results in the left and right columns are nearly the same as those in the bottom-right plots of Figures 5 and 6, respectively.

B. Experiments

The calibration methods were evaluated on real images for four test setups at three locations. Setups 1 and 2 were captured by a permanently mounted PTZ camera, which provided the ground truth tilt angle using its internal encoders. Setups 3 and 4 were captured by a portable camera without the capability of providing ground truth tilt. For both cameras, the ground truth focal lengths were computed using the planar calibration method of Zhang [23]. All other ground truth values were acquired using a tape measure, and all image measurements were obtained manually by clicking on the image to avoid image processing errors. We would like to emphasize that the distance measurements (ω) used in the overconstrained (OC) solution were different from the test distances used to evaluate the calibration accuracy.

The results of the various methods are presented in Table II. In addition to the camera parameters, the table shows the

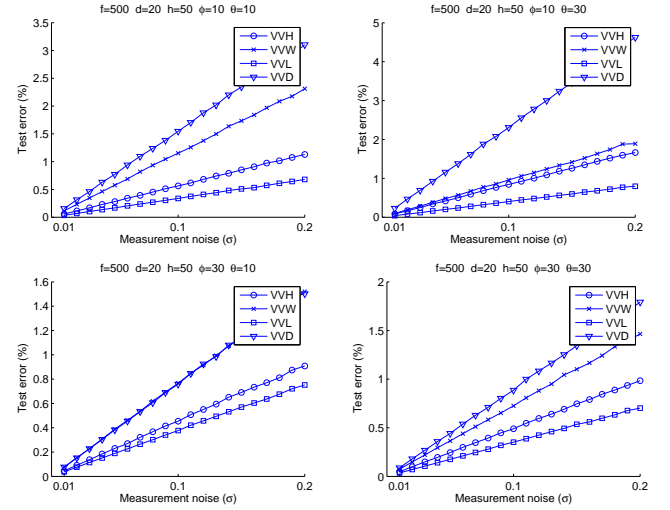


Fig. 5. Simulation results of the two-vanishing-point methods. VVL is the least sensitive to measurement noise.

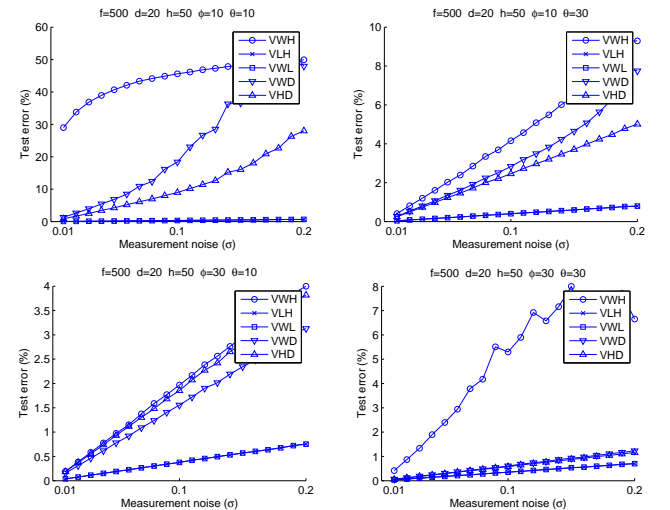


Fig. 6. Simulation results of the single-vanishing-point methods. VWL and VLH are the least sensitive to measurement noise.

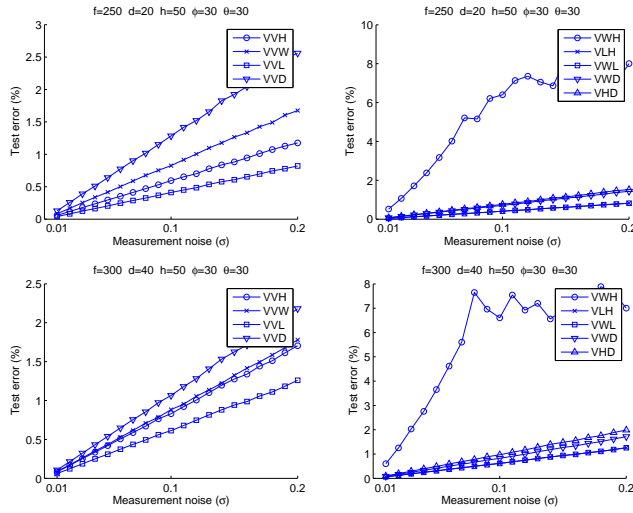


Fig. 7. Simulation results for the two-vanishing-point methods (left column) and the single-vanishing-point methods (right column) for different values of f and $\mu = h/d$. Note that the sensitivity of the methods is not significantly affected.

accuracy of each method's ability to estimate test distance measurements in the road. The test errors are then graphically illustrated in the plots of Figure 8. These results show that methods based on using two vanishing points (VVH, VVW, VVL, and VVD) and/or using a measurement of known length (VVL, VLH, and VWL) outperformed the other methods, producing errors less than 10%. In contrast, the single-vanishing point methods without a known length (VWH, VWD, and VHD) performed poorly, producing errors of approximately 10–20%. Note that even though VVD performs comparably to the single-vanishing point techniques, it is noticeably worse than any other two-vanishing point technique, thus showing that distance to the road is not a reliable measurement for calibration. These observations are consistent with the simulations above.

These results were obtained by first compensating for the non-zero roll angle in Setups 3 and 4, which was achieved by a simple image rotation. Regarding the other assumptions in the camera model, the unity aspect ratio and zero skew angle assumption hold true for nearly all modern cameras and therefore incur little risk. While the principal point can often be far from the image center, we have found that this assumption nevertheless does not significantly affect the results, causing errors to increase by no more than approximately 2%.

We should mention that while performing calibration with real images, the quadratic equation in m occasionally resulted in two positive solutions (resulting in two sets of parameters satisfying all respective equations). These cases are indicated by an asterisk (*) next to the estimated focal length in Table II. To resolve this ambiguity, we always selected m_+ as the solution when this occurred. This choice was based on the assumption that typically both the pan and tilt angles are not excessively large. Within the range $0^\circ < \theta < 45^\circ$, m_+ is guaranteed to be the correct solution for VWL, as seen in Figure 3. Similarly, if both ϕ and θ are no greater than 38.17° , then m_+ is also guaranteed to be the correct solution for VLH;

this value comes from intersecting the $b^2 = 4ac$ (gray/white) curve with the $\phi = \theta$ line. With VWH and VHD, there is only a slight chance of m_+ being the wrong solution (see the white sliver in Figures 3 and 4).

To test the validity of our choice of m_+ to overcome the ambiguity when both roots are positive, we used additional information about known quantities collected for these four setups. In the case of VWH, VLH, and VWL we used the known distance d to the road, whereas in the case of VHD we used a known length l . In all cases we found that the parameter value computed using m_+ was indeed closer to the true value than the alternative value computed using m_- . While this is not a guarantee that the “select m_+ when in doubt” heuristic will work in all situations, it nevertheless confirms its correctness in these setups, and it indicates that this heuristic may be useful in other configurations as well.

VIII. CONCLUSION

As vision-based traffic monitoring and data collection become more widespread, it will be important to accurately calibrate roadside cameras to compute vehicle speeds and classes. Generic camera calibration techniques do not readily apply to this scenario due to the lack of a calibration grid and the dominance of a single plane. While many researchers have addressed the problem of camera calibration in the context of traffic applications, there has been no attempt to date to evaluate the different approaches. In this paper we have proposed a taxonomy of roadside calibration approaches to better facilitate discussion of the relative merits of the approaches, as well as the conditions under which they are expected to work. In developing this taxonomy, we have introduced several new methods that have not been considered in previous work (VVH, VVL, VLH, VVD, VWD, and VHD), in addition to three that have (VVW, VWH, and VWL). We also introduced an overconstrained approach that not only reduces error but also overcomes the important but oft-neglected ambiguity that is inherent with single-vanishing point solutions. We then compared the methods using both simulations and real image data to evaluate their accuracy under different scenarios. In general, the methods that use a known length outperform the other methods, and all methods perform better when the camera is tilted at least a moderate amount. One important conclusion of this study is that the most popular method to date, VVW, does not perform well at estimating lengths in the road.

In this paper we have focused solely upon the geometric relationships between the various quantities, without addressing the vast number of image processing techniques for automatically obtaining them. Our purpose in doing this was to determine the efficacy of roadside camera calibration itself without regard to the errors of any particular image processing algorithm. Even with very accurate measurements, we conclude that camera calibration is a delicate and sensitive process for which great care must be taken to ensure accurate results.

| Setup | Quantity | GT | VVH | VVW | VVL | VVD | VWH | VLH | VWL | VWD | VHD | OC |
|---------|--------------------|------|--------------|--------------|--------------|--------------|-------------|-------------|---------------|---------------|---------------|-------------|
| Setup 1 | f (pixels) | 435 | 373.8 | 373.8 | 373.8 | 373.8 | 159.6 | 378.0* | 348.9 | 182.4* | 204.7* | 319.1 |
| | h (feet) | 28.5 | 28.5 | 30.5 | 28.8 | 25.4 | 28.5 | 28.5 | 30.5 | 29.6 | 28.5 | 28.7 |
| | d (feet) | 20.5 | 22.3 | 23.9 | 22.6 | 20.5 | 18.0 | 22.4 | 23.7 | 20.5 | 20.5 | 20.5 |
| | ϕ (degrees) | 14.5 | 13.9 | 13.9 | 13.9 | 13.9 | 30.1 | 13.8 | 14.9 | 26.9 | 24.4 | 14.7 |
| | θ (degrees) | — | 12.4 | 12.4 | 12.4 | 12.4 | 24.7 | 12.3 | 13.2 | 22.5 | 20.7 | 13.2 |
| | Test error | — | 6.9% | 6.8% | 6.5% | 14.6% | 32.2% | 6.8% | 5.1% | 28.3% | 28.2% | 3.5% |
| Setup 2 | f (pixels) | 435 | 320.8 | 320.8 | 320.8 | 320.8 | 324.0 | 347.2* | 312.6 | 243.8* | 257.8* | 312.6 |
| | h (feet) | 28.5 | 28.5 | 30.0 | 29.6 | 26.0 | 28.5 | 28.5 | 29.9 | 29.3 | 28.5 | 29.9 |
| | d (feet) | 20.5 | 22.8 | 23.9 | 23.7 | 20.5 | 22.8 | 23.4 | 23.7 | 20.5 | 20.5 | 20.5 |
| | ϕ (degrees) | 24 | 26.4 | 26.4 | 26.4 | 26.4 | 26.1 | 24.6 | 26.9 | 33.1 | 31.7 | 26.9 |
| | θ (degrees) | — | 24.6 | 24.6 | 24.6 | 24.6 | 24.4 | 23.2 | 25.0 | 29.4 | 28.4 | 25.0 |
| | Test error | — | 4.8% | 9.8% | 8.4% | 7.0% | 5.0% | 6.3% | 9.3% | 6.2% | 3.0% | 4.7% |
| Setup 3 | f (pixels) | 290 | 425.5 | 425.5 | 425.5 | 425.5 | 367.1 | 366.2* | 296.5 | 268.5* | 313.8* | 360.1 |
| | h (feet) | 32 | 32.0 | 35.9 | 28.5 | 26.6 | 32.0 | 32.0 | 37.1 | 37.4 | 32.0 | 40.5 |
| | d (feet) | 10 | 14.9 | 16.7 | 13.2 | 10.0 | 13.9 | 14.0 | 14.2 | 10.0 | 10.0 | 10.0 |
| | ϕ (degrees) | — | 12.0 | 12.0 | 12.0 | 12.0 | 13.9 | 13.9 | 17.0 | 18.7 | 16.1 | 16.5 |
| | θ (degrees) | — | 19.9 | 19.9 | 19.9 | 19.9 | 22.5 | 22.6 | 26.9 | 29.0 | 25.7 | 26.9 |
| | Test error | — | 11.2% | 14.6% | 10.6% | 15.6% | 5.9% | 6.0% | 3.1% | 5.1% | 10.1% | 2.8% |
| Setup 4 | f (pixels) | 290 | 314.4 | 314.4 | 314.4 | 314.4 | 705.5 | 355.6* | 281.8* | 226.9* | 293.2* | 325.9 |
| | h (feet) | 32 | 32.0 | 36.6 | 34.8 | 30.7 | 32.0 | 32.0 | 37.2 | 38.9 | 32.0 | 39.7 |
| | d (feet) | 20 | 26.5 | 30.3 | 28.9 | 20.0 | 32.2 | 27.9 | 29.2 | 20.0 | 20.0 | 20.0 |
| | ϕ (degrees) | — | 14.5 | 14.5 | 14.5 | 14.5 | 6.5 | 12.8 | 16.0 | 19.7 | 15.5 | 16.2 |
| | θ (degrees) | — | 29.3 | 29.3 | 29.3 | 29.3 | 14.4 | 26.5 | 31.9 | 37.1 | 30.9 | 31.9 |
| | Test error | — | 7.4% | 6.0% | 4.1% | 10.9% | 38.4% | 6.4% | 3.0% | 4.9% | 10.3% | 2.5% |

TABLE II

EXPERIMENTAL RESULTS ON REAL IMAGES OBTAINED WITH FOUR SETUPS (TWO DIFFERENT CAMERAS, THREE DIFFERENT LOCATIONS, AND VARIOUS ANGLES). THE COLUMN GT CONTAINS GROUND TRUTH VALUES WHEN AVAILABLE. THE REMAINING COLUMNS CONTAIN THE VALUES ESTIMATED BY THE VARIOUS METHODS, ALONG WITH THE MEAN PERCENTAGE ERROR OF THE TEST DISTANCES. ESTIMATED VALUES WITH LESS THAN 10% ERROR ARE SHOWN IN BOLD. THE ASTERISK (*) INDICATES THAT BOTH ROOTS WERE POSITIVE AND THEREFORE THE “SELECT m_+ WHEN IN DOUBT” HEURISTIC WAS EMPLOYED.

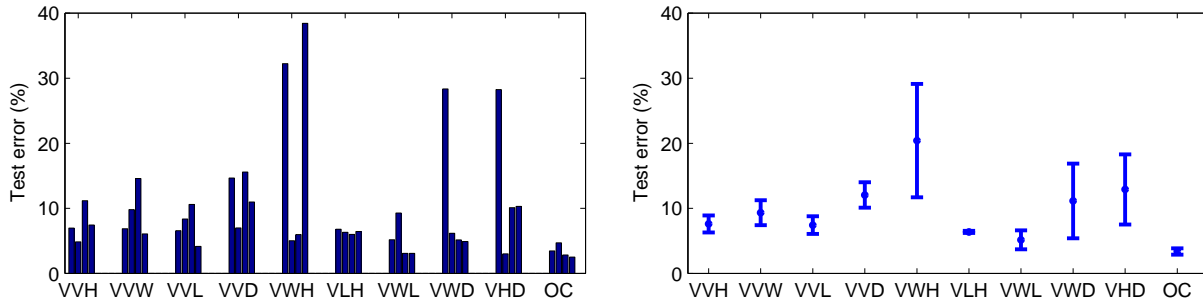


Fig. 8. Graphical illustration of the test-length errors of Table II. Left: The errors of each method for the four different setups (the four bars are for Setups 1 to 4, from left to right). Right: Mean and standard deviation of the length-measurement error among the four setups, for each method.

REFERENCES

- [1] E. K. Bas and J. D. Crisman, “An easy to install camera calibration for traffic monitoring,” in *Proceedings of the IEEE Conference on Intelligent Transportation Systems*, 1997, pp. 362–366.
- [2] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, “A real time computer vision system for measuring traffic parameters,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1997, pp. 495–501.
- [3] D. Dailey, F. W. Cathy, and S. Pumrin, “An algorithm to estimate mean traffic speed using uncalibrated cameras,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 2, pp. 98–107, June 2000.
- [4] X. Feng, M. Cao, H. Wang, and M. Collier, “The comparison of camera calibration methods based on structured-light measurement,” in *Proceedings of the 2008 Congress on Image and Signal Processing*, vol. 2, 2008, pp. 155–160.
- [5] G. S. K. Fung, N. H. C. Yung, and G. K. H. Pang, “Camera calibration from road lane markings,” *Optical Engineering*, vol. 42, pp. 2967–2977, Oct. 2003.
- [6] R. Guerchouche and F. Coldefy, “Robust camera calibration and evaluation procedure based on images rectification and 3D reconstruction,” in *Proceedings of the ICVS Workshop on Camera Calibration Methods for Computer Vision Systems (CCMVS)*, 2007, pp. 1–6.
- [7] S. Gupte, O. Masoud, R. F. K. Martin, and N. P. Papanikolopoulos, “Detection and classification of vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 1, pp. 37–47, Mar. 2002.
- [8] X. C. He and N. H. C. Yung, “New method for overcoming ill-conditioning in vanishing-point-based camera calibration,” *Optical Engineering*, vol. 46, no. 3, pp. 1–12, 2007.
- [9] F. Huang, S.-K. Wei, and R. Klette, “Comparative studies of line-based panoramic camera calibration,” in *IEEE Workshop on Omnidirectional Vision*, 2003, pp. 1–6.
- [10] N. K. Kanhere and S. T. Birchfield, “Real-time incremental segmentation and tracking of vehicles at low camera angles using stable features,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 148–160, Mar. 2008.
- [11] N. K. Kanhere, S. T. Birchfield, and W. A. Sarasua, “Automatic camera calibration using pattern detection for vision-based speed sensing,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2086, pp. 30–39, 2008.
- [12] A. H. S. Lai and N. H. C. Yung, “Lane detection by orientation and length discrimination,” *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, vol. 30, no. 4, pp. 539–548, Aug. 2000.
- [13] P. G. Michalopoulos, “Vehicle detection video through image processing: the Autoscope system,” *IEEE Transactions on Vehicular Technology*,

vol. 40, no. 1, pp. 21–29, Feb 1991.

- [14] T. N. Schoepflin and D. J. Dailey, “Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 4, no. 2, pp. 90–98, June 2003.
- [15] K.-T. Song and J.-C. Tai, “Dynamic calibration of pan-tilt-zoom cameras for traffic monitoring,” *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, vol. 36, no. 5, pp. 1091–1103, Oct. 2006.
- [16] C. Strecha, W. von Hansen, L. V. Gool, P. Fua, and U. Thoennessen, “On benchmarking camera calibration and multi-view stereo for high resolution imagery,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.
- [17] M. Trajković, “Interactive calibration of a PTZ camera for surveillance applications,” in *Asian Conference on Computer Vision*, 2002, pp. 1–8.
- [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 Journal of Robotics and Automation*, vol. RA-3, no. 4, pp. 323–344, Aug. 1987.
- [19] L. L. Wang and W. H. Tsai, “Camera calibration by vanishing lines for 3-D computer vision,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 370–376, Apr. 1991.
- [20] B.-F. Wu, W.-H. Chen, C.-W. Chang, C.-C. Liu, and C.-J. Chen, “Dynamic CCD camera calibration for traffic monitoring and vehicle applications,” in *IEEE International Conference on Systems, Man, and Cybernetics (SMC 2007)*, Oct. 2007, pp. 1717–1722.
- [21] G. Zhang, R. P. Avery, and Y. Wang, “Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1993, pp. 138–147, 2007.
- [22] Z. Zhang, M. Li, K. Huang, and T. Tan, “Practical camera auto-calibration based on object appearance and motion for traffic scene visual surveillance,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.
- [23] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [24] C. Zhaoxue and S. Pengfei, “Efficient method for camera calibration in traffic scenes,” *Electronics Letters*, vol. 40, no. 6, pp. 368–369, Mar. 2004.
- [25] H. Zollner and R. Sablatnig, “Comparison of methods for geometric camera calibration using planar calibration targets,” in *Proceedings of the 28th Workshop of the Austrian Association for Pattern Recognition*, 2004, pp. 237–244.



Neeraj K. Kanhere received the B.E. degree in Electronics Engineering from the University of Mumbai, Mumbai, India, in 2002, and the M.S. and Ph.D. degrees in Electrical Engineering from Clemson University, Clemson, SC, in 2005 and 2008, respectively. He is currently a postdoc in Electrical Engineering at Clemson University. For his research on vision-based traffic surveillance, he received the Dwight David Eisenhower Transportation Fellowship and the Clemson University Outstanding Performance Fellowship. He is the author or co-

author of more than ten conference and journal papers on computer vision, image processing, and pattern recognition applied to Intelligent Transportation Systems.



Stanley T. Birchfield (S’91–M’99–SM’06) received the B.S. degree in Electrical Engineering from Clemson University, Clemson, SC, in 1993, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1996 and 1999, respectively. While at Stanford, his research was supported by a National Science Foundation Graduate Research Fellowship, and he was part of the winning team of the AAAI Mobile Robotics Competition of 1994. From 1999 to 2003 he was a research engineer with Quindi Corporation, a startup company in Palo Alto, CA.

Since 2003 he has been with the Electrical and Computer Engineering Department at Clemson University, where he is now an Associate Professor. His research interests include visual correspondence, tracking, and segmentation, particularly applied to real-time systems.