A new constraint on the imaged absolute conic from aspect ratio and its application

Yihong Wu *, Zhanyi Hu

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, No. 95 EastRoad of Zhong Guan Cun, P.O. Box 2728, Beijing 100080, PR China

Received 1 November 2004
Available online 15 December 2004

Communicated by E. Backer

Abstract

Camera self-calibration is an important task in computer vision. In the literature, if the aspect ratio is known, people need additionally zero-skew-assumption to generate a constraint on the image of the absolute conic for camera calibration. However usually camera skew is nonzero and unknown. In this paper, a new quadric constraint on the image of the absolute conic is introduced, which is solely from known aspect ratio. Its application to single-view-based calibration and reconstruction is reported to illustrate its applicability and usefulness. In addition, the new constraint is experimentally shown to be advantageous over the commonly used zero-skew-based constraint in terms of calibration accuracy.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Absolute conic; Camera intrinsic parameter; Single-view-based camera calibration

1. Introduction

Camera self-calibration is an important task in computer vision and has been extensively studied in recent years. The aim of self-calibration is in essence to determine the image of the absolute conic, or its dual, the image of the absolute quadric (Faugeras et al., 1992; Hartley and Zisserman, 2000; Heyden and Astrom, 1997; Luong and Faugeras, 1997; Maybank and Faugeras, 1992; Pollefeys et al., 1998, 1999; Triggs, 1997). If some prior knowledge on camera intrinsic parameters is available, usually more calibration constraints could be obtained. To this end, it is essential to
convert such prior knowledge on the intrinsic parameters to some kinds of constraints on the image of the absolute conic. In the past, some constraints on the image of the absolute conic or its duality from prior knowledge were set up (Hartley and Zisserman, 2000; Heyden and Astrom, 1997; Pollefeys et al., 1998, 1999). For example, let $C$ be the $3 \times 3$ symmetric matrix of the image of the absolute conic, if the skew is zero, $C_{12} = 0$; if the aspect ratio $r$ is known and the skew is zero, $C_{22} = C_{11} r^2$. However usually camera skew is non-zero and unknown, a question comes up: Can we have a constraint on the image of the absolute conic from only known aspect ratio? To our knowledge, no reports are yet available on this problem in the literature.

In this paper, by directly representing the camera intrinsic parameters from the image of the absolute conic or its duality, any prior knowledge on the camera intrinsic parameters are easily converted into constraints on the image of the absolute conic. In particular, we show that a quadratic constraint on the image of the absolute conic can be set up from only known aspect ratio, which will be elaborated in Section 2.

This new constraint can in principle be used in any vision problem where a camera calibration step is involved either for enhancing robustness or removing ambiguities. For examples, when the camera undergoes a planar motion (Armstrong et al., 1996; Faugeras et al., 1998); when the camera is fixed in front of objects rotated on a turntable (Fitzgibbon et al., 1998; Jiang et al., 2002, 2003); when scene contains parallel, orthogonal geometric entities, circles etc. (Criminisi et al., 1999; Criminisi, 2001; Huang et al., 2004), the proposed constraint in this work can be used to remove the ambiguities in these studies. In this work, its application to camera calibration from a single image is only reported to illustrate its validity and applicability. In addition, a comparative study with the zero-skew-based constraint is also performed, and the new constraint is shown to be advantageous in terms of calibration accuracy.

People might think that the skews of most cameras are zero, and this new constraint is insignificant. We would disagree. Based on our proper experiences, we find that in most cases, camera skew is in fact non-zero. Worse still, as we reported later, mistakenly and forcefully setting the skew of a camera to zero could bring gross calibration and reconstruction errors. Of course, this new constraint should not be over-emphasized either, it is just a quadric constraint on the image of the absolute conic.

In this work, a boldface letter denotes a matrix or a column vector, a matrix or a vector with superscript “T” denotes its transpose, and “\(\approx\)” denotes the equality up to a scalar.

2. A new constraint on the imaged absolute conic from aspect ratio

Under a pinhole camera, a point $X$ in space is projected to a point $x$ in the image by
\[ x \approx K[R, t]X, \]
where $R, t$ are a $3 \times 3$ rotation matrix and a 3-vector of translation, and
\[ K = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \]
is the matrix of camera intrinsic parameters, $f_u, f_v$ ($f_u > 0$, $f_v > 0$) are the focal lengths, $s$ the skew, and $(u_0, v_0)$ the principal point. We assume the camera model is of the pinhole one throughout this work, and possible model distortions will not be considered.

The absolute conic is a virtual conic in 3D projective space consisting of points $X = (x^T, 0)^T$ on the plane at infinity such that
\[ x^T x = 0, \]
where $x = (x_1, x_2, x_3)^T$. Let $m$ be a point on the image of the absolute conic, and its corresponding space point on the absolute conic be $P = (p^T, 0)$, then by (1), we know that $m \approx KRp$, or $p \approx R^{-1}K^{-1}m$, so by $p^T p = 0$, we obtain $m^T K^{-T} K^{-1} m = 0$. It follows that the symmetric matrix of the image of the absolute conic is $K^{-T} K^{-1}$, denoted as $C$.

The ratio $f_u/f_v$ is the aspect ratio, denoted as $r$. 

Proposition 1. The camera intrinsic parameters can be represented out by \( r \) and \( C \) as

\[
\begin{align*}
    f_u^2 &= \frac{\text{Det}(C)}{C_{11}(C_{11}C_{22} - C_{12}^2)} = \frac{\text{Det}(C)}{C_{11}r^2}, \\
    f_v^2 &= \frac{C_{11}\text{Det}(C)}{(C_{11}C_{22} - C_{12}^2)^2} = \frac{\text{Det}(C)}{C_{11}r^4}, \\
    s &= -\frac{C_{12}}{C_{11}}f_v, \\
    u_0 &= \frac{C_{12}C_{23} - C_{13}C_{22}}{C_{11}C_{22} - C_{12}^2} = \frac{C_{12}C_{23} - C_{13}C_{22}}{C_{11}r^2}, \\
    v_0 &= \frac{C_{12}C_{13} - C_{11}C_{23}}{C_{11}C_{22} - C_{12}^2} = \frac{C_{12}C_{13} - C_{11}C_{23}}{C_{11}r^2}.
\end{align*}
\]

Substituting \( K = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \end{bmatrix} \) into \( C \approx K^{-T}K^{-1} \) gives

\[
C \approx \begin{bmatrix}
    \frac{1}{f_u^2} & -\frac{s}{f_u f_v} & \frac{s u_0 - f_v v_0}{f_u^2 f_v} \\
    -\frac{s}{f_u f_v} & \frac{1}{f_v^2} + \frac{s}{f_u^2} & -\frac{s (s u_0 - f_v v_0) - v_0}{f_u^2 f_v} + \frac{s^2}{f_v^2} + \frac{1}{f_v^2}
\end{bmatrix}.
\]

Then, Proposition 1 can be directly verified by this representation of \( C \).

If there is any prior knowledge on the camera intrinsic parameters \( K \), the corresponding constraint on \( C \) can be obtained straightforwardly by Proposition 1. The following is a new constraint from only aspect ratio.

Proposition 2. If the aspect ratio \( r \) is known, we have a quadratic constraint on \( C \) as

\[
\frac{C_{22}}{C_{11}} - \left( \frac{C_{12}}{C_{11}} \right)^2 = r^2.
\]

Worth noting that in the past, the commonly used constraint from the aspect ratio is \( C_{22} = C_{11}r^2 \), which is valid only if the skew is zero (\( s = 0 \)). However, the constraint in (2) does not need any additionally information but aspect ratio. This constraint seems original, and does not appear in other places. Besides, by the duality of \( C \) and the image of the absolute quadric, we can also obtain a fourth degree constraint on the image of the absolute quadric from only known aspect ratio. However, due to its higher degree, we only consider the constraint (2) in the next section.

3. Applications

In recent years, single-view-based metrology has attracted a lot of attention (Criminisi et al., 1999; Criminisi, 2001; Huang et al., 2004), where some scene information should be known beforehand. If the scene information is not sufficient to recover the Euclidean measurements from a single view, some prior knowledge on the camera intrinsic parameters should additionally be employed. In such cases, the new constraint (2) presents itself.

In addition, the new constraint (2) can also be used to remove ambiguities in the 3D reconstruction based on turntable (Fitzgibbon et al., 1998; Jiang et al., 2002, 2003).

Here single-view-based camera calibration will be considered. Similar to the single-view-based metrology, some scene information must be available. The commonly used scene information is of parallelism, orthogonality, circles, or known lengths, known angles. From such pieces of geometric information of scene, usually the images of a pair of circular points, denoted as \( m_i, \bar{m}_i \), and the vanishing point of the orthogonal direction of the circular points, denoted as \( v \), can be computed out. In the following, we assume that \( m_i, \bar{m}_i, v \) are already computed out from the scene, and will be used in combination with the constraint (2) for calibrating a camera from a single view.

As proved in (Wu et al., 2003) and illustrated in Fig. 1, \( m_i, \bar{m}_i, v \) can produce four independent constraints on \( C \), and especially \( C \) can be explicitly expressed by \( m_i, \bar{m}_i, v \) up to one free parameter \( \lambda \):

\[
C \approx \lambda (m_i \times \bar{m}_i)(m_i \times m_i)^T - (m_i \times v)(m_i \times v)^T - (\bar{m}_i \times v)(m_i \times v)^T = 2\lambda (a \times b)(a \times b)^T + (a \times v)(a \times v)^T + (b \times v)(b \times v)^T.
\]
If some prior knowledge on the camera intrinsic parameters is available, the free parameter $k$ in (3) can be solved out, then $C$ as well as $K$ can be obtained completely. Here we assume the aspect ratio of camera is known, then constraint (2) will be used to solve out $k$.

### 4. Experiments

In this section, experiments on both simulated and real data are performed, where we assume $m_I, m_{I'}$ and $v$ can be obtained from geometric structures on the scene. In the simulated data, the aspect ratio $r$ is set known. In the real data, the used camera is a Nikon COOLPIX990, whose aspect ratio $r$ is known to be 1.

The steps of our experiments are

- **Step 1:** Extract the geometric structures from a single image.
- **Step 2:** Compute $m_I, m_{I'}$ and $v$ from the extracted geometric structures, and represent $C$ as (3).
- **Step 3:** Substitute the representation of $C$ into $(C_{22}/C_{11}) - (C_{12}/C_{11})^2 = r^2$ and solve out $\lambda$. There are two solutions, denoted as $\lambda_1, \lambda_2$. Then by substituting $\lambda_1, \lambda_2$ into the representation of $C$, two solutions of $C$ are obtained, denoted as $C_1, C_2$. Discard the one that is not positive definite.
- **Step 4:** Represent $K$ from $C_i$ by Proposition 1. If we have only one $K$, then it is the desired one. Otherwise, choose the one such that $f_u > 0, f_v > 0, f_u > |s|$, and $(u_0, v_0)$ are close to the image center.

#### 4.1. Simulated experiments

The simulated intrinsic parameters are

$$K = \begin{bmatrix} f & s & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1000 & 1 & 517 \\ 0 & 1000 & 384 \\ 0 & 0 & 1 \end{bmatrix},$$

where the aspect ratio is 1. An image of size $1000 \times 700$ pixels as shown in Fig. 2 is generated from space points on a rectangle, and two lines orthogonal to this rectangle. Gaussian noise with mean 0 and standard deviation ranging from 0 to 5 pixels is added to the image points, then the intrinsic parameters are computed by the above listed steps. For each noise level, we performed 100 independent experiments, and the averaged results are shown in Table 1. We also compute the deviations of the estimated intrinsic parameters under different noise levels, and plot them in Fig. 3. As it can be seen, with noise increasing, the standard deviations of the estimated intrinsic parameters increase, and the accuracies of the estimated intrinsic parameters decrease.

![Fig. 2. The used simulated image under aspect ratio 1, where $L_1, L_2, L_3, L_4$ are from the sides of a rectangle, and $L_5, L_6$ are from two lines orthogonal to this rectangle.](image-url)
The performance of the estimated intrinsic parameters with constraint (2) is also assessed with varying the aspect ratio as follows.

The simulated camera intrinsic parameters are

\[
K = \begin{bmatrix}
    r^*f & s & u_0 \\
    0 & f & v_0 \\
    0 & 0 & 1
\end{bmatrix} = \begin{bmatrix}
    r^*1000 & 1 & 517 \\
    0 & 1000 & 384 \\
    0 & 0 & 1
\end{bmatrix},
\]

where the aspect ratio \( r \) varies from 1 to 1.5 with a step of 0.1. The simulated space data and camera motion parameters are the same as before. Under each aspect ratio, we estimate the intrinsic parameters with noise level of 2 pixels. One hundred independent trials are done, and the averaged results are shown in Table 2. The standard deviations of the estimated intrinsic parameters under different aspect ratios are also plotted in Fig. 4. We can see that for all the estimated intrinsic parameters, the deviations are larger with \( r = 1.1 \), and then decrease with \( r \) varying from 1.1 to 1.4. For \( f, u_0, v_0 \), the deviations increase with \( r \) varying from 1.4 to 1.5.

4.2. Real experiments

Fig. 5 shows a single real image of a water cup. There are two parallel circles at the brim of this cup. The contour of this cup in the image consists of two line segments and two imaged conics from the two parallel circles. We extract this contour by Canny edge detector. Then, \( v \) is computed out from the two line segments, and \( m_1, m_1 \) are computed out from the two imaged conics (Wu et al., 2004).

Following the above listed steps, the estimated intrinsic parameters are

<table>
<thead>
<tr>
<th>Noise level (pixel)</th>
<th>( f )</th>
<th>( s )</th>
<th>( u_0 )</th>
<th>( v_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1000.000</td>
<td>1.000</td>
<td>517.000</td>
<td>384.000</td>
</tr>
<tr>
<td>1</td>
<td>998.765</td>
<td>0.786</td>
<td>520.599</td>
<td>386.687</td>
</tr>
<tr>
<td>2</td>
<td>994.256</td>
<td>–0.027</td>
<td>530.149</td>
<td>393.045</td>
</tr>
<tr>
<td>3</td>
<td>985.898</td>
<td>–1.974</td>
<td>545.382</td>
<td>403.361</td>
</tr>
<tr>
<td>4</td>
<td>973.496</td>
<td>–4.876</td>
<td>567.533</td>
<td>418.546</td>
</tr>
<tr>
<td>5</td>
<td>959.804</td>
<td>–10.985</td>
<td>596.784</td>
<td>443.497</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aspect ratio</th>
<th>( f )</th>
<th>( s )</th>
<th>( u_0 )</th>
<th>( v_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>994.256</td>
<td>–0.027</td>
<td>530.149</td>
<td>393.045</td>
</tr>
<tr>
<td>1.1</td>
<td>1038.817</td>
<td>–0.098</td>
<td>466.508</td>
<td>367.638</td>
</tr>
<tr>
<td>1.2</td>
<td>1007.576</td>
<td>0.387</td>
<td>510.405</td>
<td>383.457</td>
</tr>
<tr>
<td>1.3</td>
<td>1006.715</td>
<td>0.349</td>
<td>515.358</td>
<td>386.876</td>
</tr>
<tr>
<td>1.4</td>
<td>1005.303</td>
<td>0.131</td>
<td>516.743</td>
<td>387.386</td>
</tr>
<tr>
<td>1.5</td>
<td>1004.264</td>
<td>0.200</td>
<td>517.596</td>
<td>386.796</td>
</tr>
</tbody>
</table>
In order to verify $K_1$, we are to recover the 3D structure of the cup using $K_1$. We set up the world coordinate system as: the plane containing the lower circle at the brim of this cup as the $X$–$Y$ plane, the center of this lower circle as the origin, the line through the centers of the two circles at the brim of the cup as $Z$-axis. Then, we reconstruct the cup by the estimated $K_1$ under this world coordinate system. The reconstructed two circles are

\[
\text{Cir}_1 : \begin{cases} 
X^2 + Y^2 + 0.64 \times 10^{-16}XY + 0.11 \times 10^{-14}X \\
+0.72 \times 10^{-15}Y - 0.5536 = 0, \\
Z = 0, 
\end{cases} 
\]

\[
\text{Cir}_2 : \begin{cases} 
X^2 + Y^2 + 0.12 \times 10^{-15}XY + 0.0899X \\
+0.0787Y - 0.5394 = 0, \\
Z = 1.6917. 
\end{cases} 
\]

We then compute the ratios of the height of the recovered cup to the radii of the two recovered circles, the results are

\[
\text{ratio}_{hr_1} = 2.2737, \quad \text{ratio}_{hr_2} = 2.2958. \quad (4)
\]

They are both close to the ground truth $2.3171 = 9.5/4.1$.

For comparison purpose, the calibration and reconstruction from the same image are also carried out for the following two cases:

(i) only zero-skew, i.e., $C_{12} = 0$, is used;
(ii) zero-skew plus known aspect ratio, i.e., $C_{12} = 0$ and $C_{22} = C_{11} r^2$, are used.

Case (i) is to show whether the COOLPIX990 camera’s skew is indeed close to zero; Case (ii) is to compare the performance of our new constraint vs. the commonly used constraint in the literature.

For Case (i), since the obtained principle point is out of the image region and the reconstruction has much bigger errors, it is clearly unreasonable, and will not be further reported. It implies that the camera skew is not zero.

For Case (ii), the obtained calibration and reconstruction results are

\[
K_2 = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2887.38 & -8.25 & 517.75 \\ 0 & 2887.38 & 194.75 \\ 0 & 0 & 1 \end{bmatrix},
\]

\[
\text{Cir}_1 : \begin{cases} 
X^2 + 0.9909Y^2 + 0.18 \times 10^{-8}XY + 0.54 \times 10^{-8}X \\
-0.16 \times 10^{-7}Y - 5.4251 = 0, \\
Z = 0, 
\end{cases} 
\]

\[
\text{Cir}_2 : \begin{cases} 
X^2 + 0.9909Y^2 - 0.41 \times 10^{-8}XY + 0.3959X \\
+0.3641Y - 5.2123 = 0, \\
Z = 5.2242, 
\end{cases} 
\]

\[
\text{ratio}_{hr_1} = 2.2429, \quad \text{ratio}_{hr_2} = 2.2883. \quad (5)
\]

From (4) and (5), we can see that the results from $K_1$ are more closer to the ground truth $2.3171 = 9.5/4.1$ than those from $K_2$. In addition, with our established world coordinate system, the
equations of $\text{Cir}_1, \text{Cir}_2$ should be of the form: $X^2 + Y^2 - e_1 = 0, Z = e_2$, where $e_1, e_2$ are two scalars. Hence once again, we can conclude that the recovered $\text{Cir}_1, \text{Cir}_2$ from $K_1$ have higher accuracy than those from $K_2$.

Here is another single-view-based example. Fig. 6 is the image of a corner of our laboratory. The lines on the floor and the door are extracted by Canny edge detector, then from them, $m_i, \overline{m}_i, v$ are computed. Following the above experimental steps, the estimated intrinsic parameters are

$$K_1 = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2469.43 & 14.94 & 1177.11 \\ 0 & 2469.43 & 636.05 \\ 0 & 0 & 1 \end{bmatrix}.$$ 

With the estimated $K_1$, six points shown in black dots in Fig. 6 are reconstructed. Then $L_{ij}$, length between the reconstruction points of $m_i$ and $m_j$; $\alpha$: the average of the reconstructed angles between the vector $m_i, m_3$ and vectors $m_h, m_k$, $h \neq k \in \{0, 1, 2, 3\}$; and $\beta$: the reconstructed angle of $\angle m_i, m_0, m_3$, are computed as

$L_{01} = 60.02 \text{ cm}, \quad L_{02} = 84.92 \text{ cm}, \quad L_{03} = 120.44 \text{ cm}, \quad L_{12} = 60.03 \text{ cm}, \quad \alpha = 89.83^\circ, \quad \beta = 89.91^\circ.$

By $C_{12} = 0$ and $C_{22} = C_{11}$, the estimated intrinsic parameters $K_2$ and the corresponding measurements are

$$K_2 = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2579.90 & 0 & 1252.68 \\ 0 & 2579.90 & 684.44 \\ 0 & 0 & 1 \end{bmatrix}.$$ 

$L_{01} = 60.02 \text{ cm}, \quad L_{02} = 87.35 \text{ cm}, \quad L_{03} = 127.25 \text{ cm}, \quad L_{12} = 63.43 \text{ cm}, \quad \alpha = 89.83^\circ, \quad \beta = 89.91^\circ.$

The ground truths of the above measurements are

$L_{01} = 60 \text{ cm}, \quad L_{02} = 84.85 \text{ cm}, \quad L_{03} = 120 \text{ cm}, \quad L_{12} = 60 \text{ cm}, \quad \alpha = 90^\circ, \quad \beta = 90^\circ.$

From the above results, we can see that the errors of the estimations $L_{02}, L_{03}, L_{12}$ by $K_2$ are much bigger than those by $K_1$. So, the new constraint (2) produces much higher accuracy of 3D reconstruction than zero-skew-based constraint does. It once again shows that by carelessly assuming zero-skew, using constraints $C_{12} = 0$ and $C_{22} = C_{11}$ for calibration and reconstruction is not a safe pass.

5. Conclusions

A new quadric constraint on the image of the absolute conic is introduced. This new constraint is only from the knowledge of the aspect ratio, contrary to the previous results where additional assumption of zero-skew should be made. This new constraint could be used in any vision problem where a camera calibration step is involved for either enhancing robustness or removing possible ambiguities. Some simulations and experiments on real data are carried out to illustrate its validity and applicability. In addition, a preliminary performance comparison of the new constraint to the previous commonly used constraint in the
literature is performed on real data, and it is shown that the new constraint can produce more accurate calibration and reconstruction results.

Acknowledgments

This work was supported by the National Key Basic Research and Development Program (973) under grant no. 2002CB312104 and the Natural Science Foundation of China under grant no. 60303021.

References


