Monocular vision based pose measurement of 3D object from circle and line features

Weixin Fu, Baofu Li*
School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China

*Corresponding author’s e-mail: libf6688@163.com

Abstract. An approach for estimating object pose from establishing equations between the features and their projection is proposed. In the case that the focal length of camera is unknown, this paper establishes the mathematical model between the circular feature of 3D Object and its projection through a series of basic changes, and adds common linear feature to provide spatial geometric constraint for pose measurement, then provides the nonlinear equations with seven unknown parameters. The equations are transformed into objective function, then the algorithm of the One-Dimensional search is applied to obtain the optimal value of the focal length. The other parameters can be calculated by the focal length. Experimental results verify the validity and accuracy of the method.

1. Introduction
Monocular vision measurement system is a hot field of measurement system at present. Compared with binocular vision, the single vision has the advantage of simple structure and low cost [1]. The circle has obvious advantages of in recognition and anti-occlusion compared with point and straight line features in the single vision system and it is also the basic geometry of the object. Therefore, the use of circle feature to estimate pose is a commonly method compared with point or line feature. Under perspective geometry, a projected circle appears as an ellipse in the imaging plane, and the pose of the circle can be extracted from single imaging using the inverse projection model of the calibrated camera [2]. The literatures [3] and [4] proposed a method to measure the pose under the conditions of known and unknown circular radius. However, two possible pose solutions can be recovered under normal condition when employing a single circle. Then literature [5] used angle invariance of Euclid space to determine the real solution through two coplanar lines. Wang used two parallel lines in the space to determine the circular pose [6]. The literatures [7, 8] proposed to add the point feature for pose measurement. Zhang used the three-line configuration method to eliminated the ambiguity [9]. The literatures [10, 11] proposed the use of two or more coplanar circle for pose measurement and camera calibration. Zhang Lei installed a laser range finder to obtain depth information for pose calculation [12]. In summary, there are two limitations in the measurement based on a single circle:

- The spatial circle has rotational symmetry around its normal, and determine the pose of five dimensions at most;
- The pose measurement obtained by using the image ellipse is ambiguous.
This paper presents a new method for measurement pose of the coplanar circle. The linear constraint is added on the known radius of the target, and then the mathematical model is established based on the relationship between the features and their projections. The nonlinear equations are solved by the analytical method. Finally, the algorithm in this paper is verified in the experimental part.

2. Detection principle

It is well known that a rigid body has six degrees of freedom, indicating that the pose of the object requires six independent parameters. Nowadays modern electronic devices usually zoom automatically when they shoot video in order to obtain good shooting effect, and the focal length cannot be determined as a constant. Therefore, the focal length $f$ is actually an unknown value, which is solved simultaneously with six variables of pose. Circles and lines exist in a large number of natural scenes and artificial objects, so these features can be selected as the main features of vision inspection. For objects with clear outline, the ellipse projected as a circular feature is likely to obtain a more accurate ellipse from the image. Establish five independent equations for the above seven pose and focal length parameters. In a plane, both the line and the point have two degrees of freedom. Theoretically, a circle feature with a line or a point both can determine pose of the object. However, there is no independent point on the object strictly, and the point often exists in the form of line-line, line-plane or multi-face intersection, which is difficult to detect from image. Line features generally abound as the edges of object. The projection of the long prism is also composed of a large number of pixels. It is easy to get a line by linear fitting and it has strong noise immunity. Therefore, line has higher detection accuracy than point. In this paper, the straight line is perpendicular to the plane of the circle and does not pass through the center of the circle.

The “circle + line” feature imaging model is shown in Fig.1. the camera frame $X_c - Y_c - Z_c$ is a 3-dimensional frame with the origin as the projection center and has its $Z_c$-axis pointing to the viewing direction and passing through the origin of the image plane coordinate system. The image frame $U\text{-}V$ is located at $Z_c = f$, $f$ is the focal length. The object frame $O_0X_0Y_0Z_0$ is fixed to the measured object. The spatial circle is imaged as an ellipse on the image plane by the central projection. The line $L$ is perpendicular to the plane of the circle and its projection is $L'$ on the image plane. The plane $\pi$ in which the $L$ is can be obtained by $O_c$ and $L'$. The detection of object pose is to determine the relative pose of object frame in the camera frame according to the projection of features.

![Figure 1. The “circle + line” feature imaging model](image1.png)

![Figure 2. 3D Object motion process](image2.png)

2.1. Pose description of the rigid body

The pose of the target object can be described by the object frame $O_0X_0Y_0Z_0$. For simplicity, the object frame is established according to the following rules: the circular plane is located on the plane $O_0X_0Y_0$, the center is coincides with the origin, the line feature is perpendicular the $O_0X_0$ and intersects the $X_0$ axis. the object frame can be expressed as the camera frame $O_cX_cY_cZ_c$ through four geometric transformations: rotating $\theta$ around the $Z$-axis, rotating $\beta$ around the $X$-axis, and then rotating $\alpha$ around the $Z$-axis, and the origin of the object frame is translated to point $C(t_x, t_y, t_z)$ in Fig.2. Therefore, the pose of the object can be represented by the parameters $(\alpha, \beta, \theta, t_x, t_y, t_z)$.
Q is any point on the circle. Its coordinate in the object frame and the camera frame are $(x_O, y_O, z_O), (x_c, y_c, z_c)$ respectively. The transformation of Q from the object frame to the camera frame can be expressed as

$$
(x_c, y_c, z_c)^T = R(x_O, y_O, z_O)^T + (t_x, t_y, t_z)^T
$$

Where $R = \begin{bmatrix}
\cos \alpha \cos \theta - \sin \alpha \cos \beta \sin \theta & -\cos \alpha \cos \beta - \sin \alpha \cos \beta \cos \theta & \sin \alpha \sin \beta \\
\sin \alpha \cos \theta + \sin \alpha \cos \beta \sin \theta & \sin \alpha \cos \beta - \cos \alpha \cos \beta \cos \theta & -\cos \alpha \sin \beta \\
\sin \beta \sin \theta & \sin \beta \cos \theta & \cos \beta 
\end{bmatrix}$.

### 2.2. Camera model

The camera model is derived from the 3D space to the central projection of the image plane, commonly referred to as the pin-hole model. In the Fig 1, the image frame $O_iUV$ is an coordinate system in units of pixels with the $U$ and $V$ parallel to the $X_c$ and $Y_c$ and in the same direction.

If the projection of a point $P[x_c, y_c, z_c]$ in the camera frame is $p(x', y')$, the physical size of the pixel is $\mu$, and the position of $p$ in the image is $(u_p, v_p)$. For the easy calculation, this paper set $x_p = u_p - u_0, y_p = v_p - v_0$ and $(u_0, v_0)$ is the pixel coordinate of the center point of the image. $(x_p, y_p)$ is the pixel coordinate with respect to $(u_0, v_0)$, as the origin. According the formula (2), the focal length $f$ and the pixel size $\mu$ both play the proportional role in imaging, and can be combined into one quantity that is replaced by $f_0$.

$$
\begin{align*}
{x_p} &= \frac{x'}{\mu} = \frac{f_x}{\mu} = \frac{f_0 x_c}{x_c} \\
{y_p} &= \frac{y'}{\mu} = \frac{f_y}{\mu} = \frac{f_0 y_c}{y_c}
\end{align*}
$$

### 2.3. The circle feature

For any point $P[x_c, y_c, z_c]$ on a circle which radius is $r$, it can be regarded as a point $(r, 0, 0)$ rotating around $Z_O$ axis in the object frame, the equation about $\theta$ can be obtained

$$
(x_c, y_c, z_c)^T = R(r 0 0)^T + (t_x, t_y, t_z)^T
$$

The image coordinate (4) can be calculated by (2)

$$
\begin{align*}
{x_p} &= f_0 \frac{tx + r (\cos \alpha \cos \theta - \sin \alpha \cos \beta \sin \theta)}{tx + r \sin \beta \sin \theta} \\
{y_p} &= f_0 \frac{ty + r (\sin \alpha \cos \theta + \cos \alpha \cos \beta \sin \theta)}{tx + r \sin \beta \sin \theta}
\end{align*}
$$

The six parameters of the elliptic equation (5) can be obtained through the identity $\sin^2 \theta + \cos^2 \theta = 1$ to eliminate the $\theta$.

$$
A \frac{x_p}{f_p}^2 + B \frac{y_p}{f_p}^2 + C \frac{x_p}{f_p} \frac{y_p}{f_p} + D \frac{x_p}{f_p} + E \frac{y_p}{f_p} + F = 0
$$

$$
\begin{align*}
A &= - (\sin \alpha \sin \beta r^2) + (\cos \alpha t_z)^2 + (\cos \alpha \cos \beta t_z - \sin \beta t_y)^2 \\
B &= - (\cos \alpha \sin \beta r^2) + (\cos \alpha t_z)^2 + (\sin \alpha \cos \beta t_z + \sin \beta t_z)^2 \\
C &= (\sin \alpha \cos \beta \sin \beta r^2 - \sin \alpha \cos \alpha t_z^2) + (\cos \alpha \cos \beta t_z - \sin \beta t_y)(\sin \beta t_x + \sin \alpha \cos \beta t_z) \\
D &= - (\sin \alpha \sin \beta \cos \beta) (\sin \alpha t_z - \cos \alpha t_y) - (\cos \alpha \cos \beta \sin \beta t_x + \sin \beta t_y) (\cos \alpha t_x + \sin \alpha \cos \beta t_z) \\
E &= \cos \alpha \sin \beta \cos \beta \sin \beta (\cos \alpha t_x + \sin \alpha \cos \beta t_z) (\cos \alpha t_x + \sin \alpha \cos \beta t_z) \\
F &= -(\cos \beta r)^2 + (\sin \alpha t_x - \cos \alpha t_y)^2 + ((\cos \alpha t_x + \sin \alpha \cos \beta t_z) \cos \beta)^2
\end{align*}
$$
2.4. The line feature

The line $L$ passes through the points $Q_0(d, 0, 0)$ and $Q_1(d, 0, 1)$ in object frame, where $d$ is the distance between $L$ to the center of the circle. Then the coordinate of $Q_0$ in the camera frame can be obtained by (1)

$$
\begin{align*}
X_{q0} &= t_x + d(\cos \alpha \cos \theta - \sin \alpha \cos \beta \sin \theta) \\
Y_{q0} &= t_y + d(\sin \alpha \cos \theta + \cos \alpha \cos \beta \sin \theta) \\
Z_{q0} &= t_z + d\sin \beta \sin \theta
\end{align*}
$$

(7)

The coordinate of the projection point $q_0(x_{q0}, y_{q0})$ of $Q_0$ is

$$
\begin{align*}
x_{q0} &= f_0 \frac{t_x + d(\cos \alpha \sin \theta + \sin \alpha \cos \beta \cos \theta)}{t_x + d\sin \beta \sin \theta} \\
y_{q0} &= f_0 \frac{t_y + d(\sin \alpha \sin \theta - \cos \alpha \cos \beta \cos \theta)}{t_x + d\sin \beta \sin \theta}
\end{align*}
$$

(8)

Similarly, the projection of point $Q_1$ is $q_1(x_{q1}, y_{q1})$. The projection of the line is still a line. According to the $q_0$ and $q_1$, the projection linear equation can be obtained (9)

$$
mx + ny = 1
$$

(9)

$$
m = \frac{d(\cos \alpha \sin \theta + \sin \alpha \cos \beta \cos \theta) + t_y \cos \beta + t_z \cos \alpha \sin \beta}{f_0(\cos \alpha \sin \theta + \sin \alpha \cos \beta \cos \theta) + t_y \cos \beta + t_z \cos \alpha \sin \beta}
$$

(10)

$$
n = \frac{d(\sin \alpha \sin \theta - \cos \alpha \cos \beta \cos \theta) - t_x \cos \beta + t_z \sin \alpha \sin \beta}{f_0(\cos \alpha \sin \theta + \sin \alpha \cos \beta \cos \theta) + t_y \cos \beta + t_z \cos \alpha \sin \beta}
$$

(11)

Till then, a total seven equations (6) and (11) can be used to solve seven unknown parameters including pose and focal length.

2.5. Pose measurement

It takes time to solve the seven nonlinear equations by using the general numerical method [13] and the correct solution may not always be obtained. According to (6), the coefficient of the elliptic equation is independent of $\theta$. Five pose parameters $(\alpha, \beta, t_x, t_y, t_z)$ can be determined if $f_0$ is known. At present, the pose measurement method based on circular feature projection have been mature, it can be solved directly by analytic method [10].

The equation of elliptic cone in space is obtained according to the central projection model.

$$
[X_c \ y_c \ z_c]Q[X_c \ y_c \ z_c]^T = 0
$$

(12)

Where $Q = \begin{bmatrix} A f_0^2 & C f_0^2 / 2 & D f_0 / 2 \\ C f_0^2 / 2 & B f_0^2 & E f_0 / 2 \\ D f_0 / 2 & E f_0 / 2 & F \end{bmatrix}$. Since the form of (12) is complex, it is not conductive to obtain the equation of the circle plane. By observing equation (12), $Q$ is $3 \times 3$ symmetric matrix, and it must be diagonalizable by matrix $P$.

$$
P^{-1}QP = P^TQP = diag(\lambda_1 \ \lambda_2 \ \lambda_3)
$$

(13)

Let be

$$
[X_c \ y_c \ z_c]^T = P[x' \ y' \ z']^T
$$

(14)

The above elliptic cone can be rotated to a coordinate system $O'X'Y'Z'$, where it is a standard elliptic cone, the equation is

$$
\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 = 0
$$

(15)
According to the rule of the standard elliptic cone, $\lambda_1, \lambda_2, \lambda_3$ meets the criteria: (1) $\lambda_1$, $\lambda_2$ and $\lambda_3$ have different signs. (2) $|\lambda_1| > |\lambda_2|$. $P$ is adjusted to its corresponding eigenvector. In the $O'X'Y'Z'$, the $Z'$-axis is the same as the cone axis, and major axis of the resulting ellipse will be parallel to the $Y'$-axis if the cone is intersected with a plane parallel to the $O'X'Y'$ plane.

As shown in Fig.3, the $O'L_1$ and $O'L_2$ (16) are the projections of the prisms of the cone on the plane $X'O'Z'$. $p_1p_2$ is the line that a plane $\omega$ which is parallel to $Y'$-axis is projected onto plane $X'O'Z'$. Supposed that $p_1p_2$ is parallel to the $X'$-axis, intersecting $O'L_1$ and $O'L_2$ at $p_1$ and $p_2$, and $Z'$-axis at $G_1$. Set $O'G_1 = 1$. To obtain the plane of the circle is equivalent to $p_1p_2$ rotating around $G_1$, so that the $p_1'p_2'$ intersects with the cone to form a circle after cutting the plane of the cone. The rotation direction of $p_1p_2$ can be determined by the additional line $L$. Let the normal vector of the plane $\pi$ is $n_\pi$ in the camera frame in Fig.1, so the projection vector of the $n_\pi$ on the plane $X'O'Z'$ is $n'_\omega$, and $n_\omega$ is the normal vector of the plane $\omega$. Thus $n'_\omega \perp n_\omega$ is correct by the invariance of angle in Euclidean space. Therefore, the rotation direction of the $p_1p_2$ should be consistent with the direction of $\gamma$ toward $\frac{\pi}{2}$, which can avoid the ambiguity problem of the circle. The plane $\omega$ (18) can be calculated by (17), where "±" depends on the direction of rotation.

\[
  z' = \pm \sqrt{|\lambda_1|/|\lambda_3|} x' + 1
\]

\[
  \gamma = \cos^{-1}(n''_\omega \cdot n_\omega)
\]

\[
  z' = \pm \sqrt{(|\lambda_1| - |\lambda_2|)/(|\lambda_2| + |\lambda_3|)} x' + 1
\]

![Figure 3. Plane Cutting](image)

$P'_1P'_2$ is the actual projection line of the circular plane, and $G_2$ is the intercept of the circular plane, which can be obtained according to the similar triangle principle

\[
  |O'G_2| = |P'_1P'_2|/|p_1'p_2'|
\]

Where $|p_1'p_2'| = \frac{2}{|\lambda_2|} \sqrt{\frac{|\lambda_1||\lambda_2|(|\lambda_1|+|\lambda_3|)}{|\lambda_1|+|\lambda_3|}}$, and $P'_1P'_2$ can be expressed as

\[
  z' = \pm \sqrt{(|\lambda_1| - |\lambda_2|)/(|\lambda_2| + |\lambda_3|)} x' + R|A_2|\sqrt{(|\lambda_1| + |\lambda_3|)/(|\lambda_1||\lambda_2|(|\lambda_1|+|\lambda_3|))}
\]

The center of the circle is the midpoint of $P'_1P'_2$. After simple geometric operation, the normal vector and the position of the circular plane are

\[
  [x'_0 \ y'_0 \ z'_0] = R \left[ \pm \sqrt{\frac{|\lambda_3||\lambda_1|-|\lambda_2|}{|\lambda_1||\lambda_2|(|\lambda_1|+|\lambda_3|)}} \ 0 \ \sqrt{\frac{|\lambda_1||\lambda_2|+|\lambda_3|}{|\lambda_1||\lambda_2|(|\lambda_1|+|\lambda_3|)}} \right]
\]

\[
  [n'_{ox} \ n'_{oy} \ n'_{oz}] = \left[ \pm \frac{|\lambda_3|-|\lambda_2|}{\sqrt{|\lambda_1||\lambda_2|(|\lambda_1|+|\lambda_3|)}} \ 0 \ - \frac{|\lambda_2|+|\lambda_3|}{\sqrt{|\lambda_1||\lambda_2|(|\lambda_1|+|\lambda_3|)}} \right]
\]
Where \([x'_0, y'_0, z'_0]\) is the position of the center of the circle, \([n'_0x, n'_0y, n'_0z]\) is the normal vector and "±" depends on the direction of rotation. The plane \(\omega\) is transformed back to the camera frame by (14), and \((\alpha, \beta, t_x, t_y, t_z)\) can be obtained. After calculating the (10), (23) can be got.

\[
msin \alpha - n \cos \alpha = \frac{\cos \beta}{f_0 \sin \beta}
\]  
(23)

Then the focal length is

\[
f_0 = \frac{\cos \beta}{(msin \alpha - n \cos \alpha) \sin \beta}
\]  
(24)

Since the initial value \(f_0\) is not an exact value, the \(f_0\) is repeated by (24). If the absolute value of the difference between \(\gamma\) and \(\frac{\pi}{2}\) is less than the threshold, the iteration is completed. Finally, \(\theta\) is can be obtained by (25) which is deduced through (10). The obtained parameters are taken as the initial solution of the next image, and the pose of each frame image is obtained.

\[
\sin(\theta - \varphi) = \frac{(mt_y + nt_x) \cos \beta - t_z (msin \alpha - n \cos \alpha) \sin \beta}{\sqrt{(msin \alpha - n \cos \alpha)^2 + (m \cos \alpha + n \sin \alpha)^2 \cos^2 \beta}}
\]  
(25)

Where \(\varphi = \tan^{-1}(\frac{m \cos \alpha + n \sin \alpha) \cos \beta}{msin \alpha - n \cos \alpha}\). If \((msin \alpha - n \cos \alpha) = 0\), the circular plane is projected as a line on the imaging plane. However, this situation can be avoided artificially in practical project and this paper does not discuss this case.

3. Experimental results and analysis

3.1. Experimental introduction

In this paper, the self-lubricating bearing tester for fatigue is taken as an experimental equipment, as shown in Fig 4. The column of the tester is fixed with the ground. The inner ring of bearing is trapped on the column, it can slide up and down along the column. The outer ring of bearing connects with three-ear oscillating sleeve through the press ring on the end. The angle between the hydraulic servo cylinders is 120°, they drive the tested bearing to swing and slip according to the motion set by the numerical control system. Try the pose detection method proposed in this paper to detect the movement of the three-ear oscillating sleeve.

As the end press ring is a ring, the outer edge of its upper plane has a clear outline and less shielding part, it can better reflect the movement of the bearing. Therefore, this circle is selected as one of the moving targets. There is a clearly edge at the connection between an ear of the three-ear oscillating sleeve and the matrix, and the edge is perpendicular to the end press ring, so this edge is selected as the linear feature.
3.2. Video processing and results
During the experiment, the camera was fixed for shooting, and one frame of image was selected every 0.5 seconds for processing. The processed images were 31 frames in total, and the image resolution was set to $544 \times 960$. The origin of the object frame was set as the center of the circle. The line connecting the center of the circle and the intersection point of the prism passing through the circle plane was taken as the $X_O$ axes. The normal vector of the circle was $Z_O$ axes. The circle radius $r$ is 154mm, and the distance $d$ is 180mm.

The images were identified by the computer, and the fitting result was shown in Fig.5. After obtaining the elliptic equation and the linear equation, the pose of the end press ring is solved through programming. The specific process is shown in Fig.6. The result of the final calculation and actual values are shown in Fig.7.

![Figure 6. Flow chart of pose measurement](image)

![Figure 7. The result of the final calculation and actual values](image)

As can be seen from the Fig.7, it is clear that the camera at this position has a high detection accuracy for the $\alpha, \beta, t_x, t_y$, with maximum $0.03^\circ$ of $\alpha$ error, maximum $0.04^\circ$ of $\beta$ error, maximum $0.08$mm of $t_x$ error, maximum of $0.18$mm of $t_y$ error. However, the error of $\theta, t_z$ is more than other parameters. Because the large swing range of the bearing around $Z_c$-axis. Moreover, An image taken by a camera at that position is sensitive to the rotation amount around $Z_c$-axis and the displacement in the $X_c$ and $Y_c$.

4. Conclusion
This paper presents a method to measure the pose of the object by using the features of circle and line when focal length is unknown. First, establish seven nonlinear equations through the relationship between the object and its projection, then equations are converted into object function for solving. This paper gives the specific steps of the algorithm, and it ensures the accuracy of pose calculation by applying such features theoretically. The experiment demonstrates the feasibility of the method from the experimental point of view. Because there are a lot of mixed features such as lines and circles in life, this method can be applied to the pose measurement of object in practical engineering.

References
[1] Zhao Lian jun. Research on mono-vision pose measurement based on features of target[D]. Graduate School of Chinese Academy of Sciences (The Institute of Optics and Electronics), 2014.
[2] Xu G, Li X, Su J, et al. Circle recognition and parameter measurement methods with the similarity constraint on the slope curve of the tangent lines to a contour[J]. Optik, 2013, 124(16):2463-2467
[3] Shiu, Y.C., Ahmad, S.. 3D location of circular and spherical features by monocular model-based vision[P]. Systems, Man and Cybernetics, 1989. Conference Proceedings. IEEE International Conference on,1989.

[4] CHEN Z, HUANG J B. A vision-based method for the circle pose determination with a direct geometric interpretation[J]. Proceedings of IEEE Transactions on Robotics and Automation, 1999, 15(6):1135-1140.

[5] Wei Zhen zhong, Zhao Zheng, Zhang Guang jun. Solution of duality in pose estimation of single circle using Euclidean angular constraint [J]. Optics and Precision Engineering, 2010, 18(3):685-691.

[6] Wang G, Wu J, Ji Z. Single view based pose estimation from circle or parallel lines[J]. Pattern Recognition Letters, 2008, 29(7):977-985.

[7] Miao Xi kui, Zhu Feng, Ding Qing hai, et al. Monocular vision pose measurement based on docking ring component [J]. Acta Optica Sinica, 2013, 33(4):115-123.

[8] Zhang Li min, Zhu Feng, Hao Ying Ming, et al. Pose measurement based on circular feature and heteromorphic feature [J]. Acta Photonica Sinica, 2015, 44(11):89-94.

[9] Zhang Lijun, Huang Xuexiang, Feng Weichun, et al. Solution of duality in circular feature with three line configuration [J]. Acta Optic Sinica, 2016, 36(5):0515001.

[10] Chen X, Zhao Y. The projected circle centres and polar line for camera self-calibration[J]. Optik, 2015, 126(20):2565-2570

[11] Bin Huang, Yongrong Sun, Yunfeng Zhu, Zhi Xiong, Jianye Liu. Vision pose estimation from planar dual circles in a single image[J]. Optik – International Journal for Light and Electron Optics, 2016, 127(10).

[12] Zhang Lei, Xu Kejun, Zhao Rui et al.. Improvement of position and orientation measurement algorithm of monocular vision based on circle feature [J]. HeFei University of Technology, 2009, 32(11):1669-1673.

[13] Gao, Wei-Feng; LiuLiu, San-Yang; Huang, Ling-Ling. A novel artificial bee colony algorithm with Powell's method[J]. Applied Soft Computing Journal, v 13, n 9, p 3763-3775, 2013.