Non-horizontal target measurement method based on monocular vision

Jiajin Lang, Jiafa Mao and Ronghua Liang

College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, People's Republic of China

ABSTRACT
With the development of computer vision technology, target measurement methods have been widely used in robot automatic obstacle avoidance, vehicle-assisted driving and other systems. There are many target measurement technologies available, but most of them are based on binocular or trinocular vision, or based on monocular vision with other auxiliary equipment, or based on monocular vision to measure horizontal target. The first two technologies achieve precise positioning by increasing the amount of data and sacrificing processing speed, while the third only studies the measurement method of the horizontal target. To address the complexity of the multi-equipment measurement methods and the limitation of the monocular measurement methods for horizontal target, this paper proposes a novel monocular vision measurement method for the non-horizontal target. According to the principle of camera imaging, internal and external parameters of the camera and analogue-to-digital conversion principle, the imaging relationship model for measuring the relative height and target distance of non-horizontal target is deduced, and the solvability of the model is demonstrated by mathematics. The experimental results verify the correctness and feasibility of this method.

1. Introduction
In recent years, computer vision technology has been widely used in robotics (Chen, 2021; Ferro et al., 2019; Martinez-Martin & Del Pobil, 2021; Song et al., 2021), target tracking (Javanmardi & Qi, 2018), autonomous navigation (Ou et al., 2021), vehicle-assisted driving (Liu et al., 2021; Martinez-Díaz, 2021; Rani et al., 2021) and other fields application. Mobile robots and driverless vehicles use cameras to obtain a large amount of external information, which is inseparable from the support of computer vision technology. Simple and efficient acquisition of the relative height and target distance contributes to the development of computer vision technology. The existing research on machine vision measurement technologies mainly focuses on the following three aspects:

(1) Measurement technologies of binocular or trinocular vision.

Since the human is binocular, the research interest of scholars on binocular vision (Hu et al., 2020; Kong et al., 2020; Liu, 2021; Mansour et al., 2019) has always been high. Ortiz et al. (2018) established the mathematical modelling of the depth error determined by the ZED camera considering a left RGB image and the depth map. And they applied the methodology to find the mathematical models of the RMS error of the Stereolabs ZED camera for all of its resolutions of operation. However, binocular vision measurements have problems such as the accuracy and synchronization between the two lenses, which directly affect the accuracy and speed of measurements. In some special engineering applications, such as pipeline measurement (Cheng et al., 2021; Haertel et al., 2015), binocular vision has the main deficiency of limited measurement range, and many scholars have carried out research on trinocular vision (Ge et al., 2021). Shao et al. (Shao & Dong, 2019) designed a trinocular vision system based on the camera with tilt-shift lens, which can enlarge the overlapping area of the trinocular vision system. And the trifocal tensor provides a stronger constraint for feature matching, which led to the method being stable and accurate for 3D reconstruction. However, in trinocular vision, different cameras acquire different data in the same scene, typically feature matching is normally according to epipolar constraint, and there is no universal algorithm to match different scenes.

(2) Measurement technologies of monocular vision with other auxiliary equipment.

CONTACT Ronghua Liang langjiajin@foxmail.com

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
In recent years, scholars have also used monocular vision with other auxiliary equipment for measurement, such as radar (Reina et al., 2011; Yao et al., 2019), aided measuring probe (Huang & Ye, 2005), reference wall (Yuan & He, 2020) and so on. Jiehu et al. (2020) proposed an effective calibration method of 1D LRF for large-scale 3D measurement under harsh working environment of strong sunshine and complex background. During practical measurement, the pixel coordinates of laser spots can be obtained from the calibration model instead of the imaging of laser spots by the camera. Data calibration and data fusion for different equipment are key problems that need to be solved for other auxiliary equipment and visual fusion. The fusion of equipment will inevitably lead to the increase in computational complexity. To be applied in practice, the algorithm still needs to be improved the operation speed of the system.

(3) Horizontal target measurement technology of monocular vision.

Under the comprehensive consideration of the manufacturing process, cost, accuracy and other aspects, the first two technologies are limited in marketing. While monocular vision measurement has unique practical application value in the existing market due to its low cost (Shankar et al., 2018) and reliability. Mao et al. (2017) proposed a target depth measurement method based on monocular vision, which realized target depth analysis on the premise that the height of the machine vision and the target height were known. However, it is limited to the top of the horizontal target at the centre of the image. On this basis, the team continued to study the situation where the top of the horizontal target is not in the centre of the image (Mao et al., 2020), that is, the top of the horizontal target is higher or lower than the centre of the image. They derived the mathematical model and verified the correctness of the method experimentally. However, the existing monocular measurement technologies are mostly used to measure the height or distance of the horizontal target. In practical applications such as road measurement, automatic vehicle parking, and robot babysitting, the target and machine vision are not at the same level, and it is difficult to achieve non-horizontal target measurement.

In view of the above problems, a new monocular vision measurement method for non-horizontal target is proposed. The imaging relationship model is derived from the geometric relationship between relative height and target distance. According to target self-height and camera internal and external parameters, the calculation equations of relative height and target distance are obtained. And the correctness and feasibility of the method are verified by mathematical derivation and experiments. The main contributions of this paper are as follows:

1. The imaging relationship model of relative height and target distance is theoretically derived without camera calibration.
2. In solving the problems of relative height and target distance, the theory demonstrates the solvability of two equations before and after changing camera height.
3. The experiment verifies the feasibility of the theory and the accuracy of the measurement.

2. Imaging principle of horizontal target

Before understanding the principle of non-horizontal target imaging, it is necessary to understand the principle of horizontal target imaging. The required parameters are as follows: (1) camera height, which is denoted as \( h_r \) (mm); (2) target distance, which is denoted as \( m \) (mm); (3) target height, which is denoted as \( h_0 \) (mm); (4) camera focal length, which is denoted as \( f \) (mm).

The camera imaging model can be approximated as a pinhole imaging model, as shown in Figure 1. Point F indicates the camera position, and the main optical axis coincides with line segment FP (red line in Figure 1). Take the point O where the camera is perpendicular to the horizontal plane as the origin of the world coordinate system. The horizontal line along the visual direction of the camera is the Z-axis, the line perpendicular to the Z-axis on the horizontal plane is the X-axis and the segment OF is the Y-axis, then the world coordinate system is defined as OXYZ. Among them, the camera height (segment OF) is \( h_r \) (mm), there is a target AB at the distance of \( m \) (mm) from the camera (segment OA), and the target self-height is \( h_0 \) (mm).

Also set the image plane as \( \pi \), the centre point of the image plane as point P, and the length of the segment FP as the camera focal length \( f \) (mm). The target AB is projected on the image plane \( \pi \) as the segment CD after imaging by the camera. Among them, the camera height is the external parameter of the camera, and the camera focal length, photosensitive element size and image size are the camera internal parameters of the camera. The points A, B, C, D, F, O and P are all on the YOZ plane, so the three-dimensional geometry can be simplified to two-dimensional geometry, as shown in Figure 2.

From Figure 2 and the tangent theorem, the following equations can be obtained:

\[
\tan \angle OFA = \frac{m}{h_r} \tag{1}
\]
Observing from the camera point F to the image plane \( \pi \), the image CD of the target has three positional relationships with point P: point P is at the upper end of point C, point P is within CD (as shown in Figure 3) and point P is at the lower end of point D. Consider the DC direction as the positive direction (the same below).

The following equations can be obtained:

\[
\tan \angle PFC = \tan \angle PFB = \frac{|PC|}{f} \tag{3}
\]

\[
\tan \angle PFD = \tan \angle PFA = \frac{|PD|}{f} \tag{4}
\]

If the point P is within CD as shown in Figure 3, then:

\[
\angle AFB = \angle OFB - \angle OFA = \angle PFB + \angle PFA \tag{5}
\]

In the right equation of Equations (5), both sides are taken tangent at the same time, it is easy to obtain:

\[
m = \frac{h_o + \sqrt{h_o - 4K(h_r^2 - h_rh_o)}}{2K} \tag{6}
\]

where

\[
K = \frac{f|CD|}{f^2 + PC \cdot PD} \tag{7}
\]
collectively referred to as relative height, and denoted as \( h \) (mm). And the horizontal length of the target relative to the camera is referred to as target distance, denoted as \( m \) (mm). Add an unknown quantity to the principle described in Section 2 – relative height, as shown in the segment EA (green dotted line in Figure 4), and analyse the relationship between the unknown quantities \( m \) and \( h \).

According to the imaging principle of the horizontal target in Section 2, points A, B, C, D, E, F, O and P are all on the YOZ plane. The three-dimensional geometry is now simplified to two-dimensional geometry, as shown in Figure 5.

As shown in Figure 5, point P is at the lower end of point D, and the following equations can be obtained based on the tangent theorem:

\[
\tan \angle OFA = \frac{m}{h_r - h} \quad (8)
\]

\[
\tan \angle OFB = \frac{m}{h_r - h_o - h} \quad (9)
\]

\[
\angle AFB = \angle OFB - \angle OFA = \angle PFB - \angle PFA \quad (10)
\]

In the right equation of Equation (10), both sides are taken tangent at the same time. Substituting into Equations (3), Equations (4) and Equation (7), the camera imaging relationship model of non-horizontal target can be
simply described as

\[
\left( m - \frac{h_0}{2K} \right)^2 + \left[ h - \left( h_r - \frac{h_0}{2} \right) \right]^2 = \frac{h_0^2}{4K^2} + \frac{h_0^2}{4} \tag{11}
\]

where \( K \) is shown in Equation (7).

As shown in Equation (11), there are two unknowns, namely the target distance \( m \) and the relative height \( h \). A binary quadratic relational model equation cannot solve the unique solution. Therefore, a subsequent study after changing camera height is introduced.

4. Imaging relationship model and theoretical analysis after changing camera height

4.1. Imaging relationship model after changing camera height

To realize the change of the camera height \( h_r \) in practical applications, the road measurement can be carried out by adjusting the lens height, and the monocular robot can adjust the neck or limb length. The amount of camera height change is denoted as \( \Delta h_r \). Take '+-' if the camera is far from the ground; Take '−' if the camera is close to the ground. The camera position after the camera height change is point \( F_1 \), the image of the target \( AB \) on the image plane \( \pi_1 \) is \( C_1D_1 \), the point \( P_1 \) is the centre point of the image plane \( \pi_1 \), and the length of the segment \( F_1P_1 \) is the camera focal length \( f \), as shown in Figure 6.

From the imaging principle of non-horizontal target in Section 3:

\[
\tan \angle OF_1A = \frac{m}{h_r \pm \Delta h_r - h} \tag{12}
\]

\[
\tan \angle OF_1B = \frac{m}{h_r \pm \Delta h_r - h_0 - h} \tag{13}
\]

\[
\tan \angle P_1F_1C_1 = \tan \angle P_1F_1B = \frac{|P_1C_1|}{f} \tag{14}
\]

\[
\tan \angle P_1F_1D_1 = \tan \angle P_1F_1A = \frac{|P_1D_1|}{f} \tag{15}
\]
As shown in Figure 6, if the point $P_1$ is within the segment $C_1D_1$, we can obtain

$$\angle AF_1B = \angle OF_1B - \angle OF_1A = \angle P_1F_1B + \angle P_1F_1A \tag{16}$$

In the equation on the right side of Equation (16), both sides are taken tangent at the same time. After changing the camera height, the non-horizontal target imaging relationship model is as

$$\left( m - \frac{h_0}{2K_1} \right)^2 + \left[ h - \left( h_r - \frac{h_0}{2} \right) \right]^2 = \frac{h_0^2}{4K_1^2} + \frac{h_r^2}{4} \tag{17}$$

where $K_1$ is shown in Equation (7) and $K_1$ is shown in Equation (18). The solvability analysis of Equations (19) is carried out below.

### 4.2. Solvability analysis of imaging relationship model after changing camera height

In Equation (19), the target distance $m$ and the relative height $h$ are unknown quantities, and other parameters are obtained from the camera internal and external parameters, image segmentation and target self-height, that is, other parameters are known. From the expression of Equation (19), the solution of the equation is to determine the positional relationship of two circles. The
positional relationships of the two circles, and the quantitative relationships between the centre distance $d$ and the radii of the two circles ($r_1$, $r_2$) are shown in Table 1.

Take point $G$ as the centre and $r_1$ as the radius to make circle I and take point $L$ as the centre and $r_2$ as the radius to make circle II. The coordinates of the centres and the lengths of the radii of the two circles are shown as

$$
\begin{align*}
G &= \left( \frac{h_o}{2K}, h_r - \frac{h_o}{2} \right) \\
L &= \left( \frac{h_o}{2K}, h_r \pm \Delta h_r - \frac{h_o}{2} \right)
\end{align*}
$$

Now, it is necessary to prove the positional relationship of the two circles. In this case, the imaging relationship model is determined, and the change of camera height shows that there is a unique variable. That is, camera height $h_r$, target distance $m$, relative height $h$, target
height $h_0$, and camera focal length $f$ are all constants, and the camera height change amount $\Delta h_r$ is the only variable. As shown in Figure 6, the world coordinate system is brought into the $K$ value obtained by the original image segmentation, that is, Equations (3)–(4) and Equations (8)–(10) are put into the Equation (7) to obtain:

$$K = \frac{f |\vec{CD}|}{f^2 + |\vec{PC}| |\vec{PD}|} = \frac{\tan \angle AFP - \tan \angle BFP}{1 + \tan \angle AFP \tan \angle BFP} = \tan (\angle AFP - \angle BFP)$$

$$= \tan \angle AFB = \frac{h_0 m}{h_r - \left( h + \frac{h_0}{2} \right)^2} - \frac{h_0}{4}$$

Among them, $|\vec{PC}| = f \tan \angle BFP$, $|\vec{PD}| = f \tan \angle AFP$, $|\vec{CD}| = |\vec{PC}| - |\vec{PD}|$.

Substituting Equations (12)–(16) into Equation (18), $K_1$ can be obtained in the same way:

$$K_1 = \frac{h_0 m}{(h_r \pm \Delta h_r) - \left( h + \frac{h_0}{2} \right)^2} - \frac{h_0}{4}$$

The positional relationship between the two circles is determined by the relationship between the centre distance and the radii of the two circles, where the centre distance is

$$d = |GL| = \sqrt{\frac{h_0}{2K} - \frac{h_0}{2K_1}} + \Delta h_r^2$$

Substituting the radius of Equations (20) and Equation (23) into $d^2 < (r_1 + r_2)^2$, and the simplified relation is as follows:

$$\frac{2\Delta h_r^2}{h_0^2} < 1 + \frac{1}{KK_1} + \sqrt{\frac{1}{K^2} + 1} \sqrt{\frac{1}{K_1^2} + 1}$$

where, due to the real-time requirements, $\Delta h_r$ does not change much, the value of $d^2$ is an approximate value greater than zero, and $(r_1 + r_2)^2$ is greater than 1 constantly. That is, $d^2 < (r_1 + r_2)^2$ is established.

Observing the radii equations of the two circles, Equation (21) and Equation (22), it can be seen that $r_1$ is a fixed value, and $r_2$ will change with the change of camera height, so it can be known that:

$$|r_1 - r_2| = \frac{dr_2}{d \Delta h_r} \cdot \Delta h_r = \frac{\partial r_2}{\partial K_1} \cdot \frac{\partial K_1}{\partial \Delta h_r} \cdot \Delta h_r$$

$$= \pm \frac{h_0 \left[ (h_r \pm \Delta h_r) - \left( h + \frac{h_0}{2} \right) \right]}{2r_2 K_1 m} \cdot \Delta h_r$$

Substituting the parameters of obj 3 in Table 3 into Equation (23) and Equation (25), the following Figure 7 can be obtained. Where the abscissa is the camera height
change $\Delta h_r$, the blue line represents the change of the centre distance $d$ with $\Delta h_r$, and the red line is the change of $|r_1 - r_2|$ with $\Delta h_r$. Now, circle I is regarded as a fixed reference circle, while circle II changes correspondingly with the change of $\Delta h_r$. According to Figure 7 and Equations (20), the centre distance $d$ represents the distance that the centre point L of the circle II moves relative to the centre point G of the circle I. And $|r_1 - r_2|$ represents the change of the radius $r_2$ of the circle II relative to the radius $r_1$ of the circle I. Under the requirements of complete imaging of the target and real-time calculation in Table 3, $d > |r_1 - r_2|$ is established.

To sum up, $|r_1 - r_2| < d < r_1 + r_2$ is established. Combining with Table 1, it can be seen that the two circles intersect as shown in Equations (20). Draw two circles in the coordinate system with horizontal coordinate as target distance $m$ and vertical coordinate as relative height $h$, as shown in Figure 8. The centre coordinates of circles I and II and the lengths of the radii of the two circles are shown in Equation (20). The red circle in Figure 8 is circle I, and the green circle is circle II. The two circles intersect at point S and point T. The centres of the two circles are connected, and after extending, they intersect the segment ST at point Q. The vertical line is drawn through the point S, the horizontal line is drawn through the point Q and the two lines intersect at point J. The vertical line is drawn through the centre G, the horizontal line is drawn through the centre L and the two lines intersect at point N.

It can be seen from Figure 8 that $\angle LQT$ and $\angle LQS$ are right angles, and the side $LQ$ of right triangle $LQT$ and the side $GQ$ of right triangle $GQT$ are in the same straight line, so the equations can be combined:

$$
\begin{align*}
GQ^2 + TQ^2 &= r_1^2 \quad \cdots \quad 1 \\
LQ^2 + TQ^2 &= r_2^2 \quad \cdots \quad 2 \\
LQ = GL - GQ &\quad \cdots \quad 3
\end{align*}
$$

(26)

Substitute the $LQ$ of Equation (26-3) into Equation (26-2), subtract from Equation (26-1), and simplify as

$$
GQ = \frac{r_1^2 - r_2^2 - GL^2}{2GL}
$$

(27)

The coordinates of centre G and centre L are shown in Equation (20). According to the ratio of segment $GQ$ and segment $GL$, the coordinates of point Q can be known:

$$
Q = \left\{ \begin{array}{l}
m_Q = m_G - \frac{GQ}{GL}(m_L - m_G) \\
h_Q = h_G - \frac{GQ}{GL}(h_L - h_G)
\end{array} \right.
$$

(28)

In Figure 8, the extension line of segment NG and the extension line of segment JQ intersect at point W, $\angle LGN$ and $\angle QGW$ are opposite angles, $\angle WQG$ and $\angle QGW$ are complementary, $\angle WQG$ and $\angle SQJ$ are complementary, so:

$$
\angle LGN = \angle SQJ
$$

(29)

And because $\angle QJS$ and $\angle GNJ$ are right angles, and the two angles of the two triangles are correspondingly equal, then the two triangles are similar, that is:

$$
\Delta QJS \cong \Delta GNJ
$$

(30)

For the similarity of two triangles, the three sides correspond to proportion, that is:

$$
\frac{QJ}{GN} = \frac{JS}{NL} = \frac{QS}{GL}
$$

(31)

Substitute the coordinates of point J, point Q, point N into Equations (31) and get:

$$
\begin{align*}
\frac{m_Q - m_S}{h_G - h_L} &= \frac{h_Q - h_S}{m_L - m_G} = \frac{SQ}{GL} \\
\frac{m_T - m_Q}{h_G - h_L} &= \frac{h_T - h_Q}{m_L - m_G} = \frac{SQ}{GL}
\end{align*}
$$

(32)

Simplifying Equations (32), the coordinates of point S and point T can be obtained as follows:

$$
\begin{align*}
S &= \left\{ \begin{array}{l}
m_S = m_Q - \frac{SQ}{GL}(h_G - h_L) \\
h_S = h_Q - \frac{SQ}{GL}(m_L - m_G)
\end{array} \right. \\
T &= \left\{ \begin{array}{l}
m_T = m_Q + \frac{SQ}{GL}(h_G - h_L) \\
h_T = h_Q + \frac{SQ}{GL}(m_L - m_G)
\end{array} \right.
\end{align*}
$$

(33)

(34)
There are only four operations of real numbers in the above analysis process, and there is no imaginary number solution. And the intersection point S and point T of the two circles do not coincide, then $SQ \neq 0$ in Equations (33) and Equations (34). As can be seen from Equations (20), point G and point L must not coincide, then $GL \neq 0 \land h_G - h_L \neq 0 \land m_L - m_G \neq 0$ in Equations (33) and Equations (34). In summary, the two circles intersect at point S and point T, and the coordinates are respectively shown in Equations (33) and Equations (34), that is, the two curves of Equations (19) intersect at point S and point T. Since the target distance $m$ cannot be negative or close to the origin (which may result in unclear imaging or incomplete target), point S is excluded. Therefore, a unique solution to the Equation (19) can be solved, as shown in Equation (34).

5. Experimental results and analysis

To verify the correctness and feasibility of the above-mentioned monocular imaging relationship model, the relative height and target distance measurement experiments after changing camera height were carried out.

5.1. Experiment preparation

In a well-lit room, five blue targets with a fixed length of 100 mm and a spacing of 97 mm are evenly distributed on the wall, and the target closest to the ground is 316 mm from the ground. The targets are all in the same line and perpendicular to the ground. From the ground up, the five targets are marked as obj1, obj2, obj3, obj4, obj5. The meaning of multiple identical targets is to obtain multiple sets of camera height changes in one image.

5.2. Experimental results

Multiple sets of experimental images of camera I and camera II are collected respectively, and the blue targets are segmented according to the RGB three channels of colour image. The blue targets are used to distinguish them from the complicated background in the image segmentation process and improve the accuracy of target segmentation, as shown in Figure 10.
Table 3. Experimental results of camera I when the target distance is 3065 mm and the camera height is 718 mm.

| (h, m) | obj₁ | obj₂ | obj₃ | obj₄ | obj₅ |
|-------|------|------|------|------|------|
| Original target | obj₁ | (–, –) | (518.6, 3001.7) | (688.9, 2995.3) | (913.0, 3001.9) | (1110.0, 3002.0) |
| | obj₂ | (321.6, 3001.7) | (–, –) | (662.1, 2995.4) | (913.5, 3001.8) | (1110.5, 3001.9) |
| | obj₃ | (295.4, 2995.3) | (465.6, 2995.4) | (–, –) | (967.2, 2992.0) | (1137.2, 2993.5) |
| | obj₄ | (322.5, 3001.9) | (520.0, 3001.8) | (770.2, 2992.0) | (–, –) | (1111.3, 3001.6) |
| | obj₅ | (322.9, 3002.0) | (520.4, 3001.9) | (743.6, 2993.5) | (914.7, 3001.6) | (–, –) |
| Truth value | (316.0, 3065.0) | (513.0, 3065.0) | (710.0, 3065.0) | (907.0, 3065.0) | (1104.0, 3065.0) |

Figure 11. Diagram of experimental results of camera I when the target distance is 3065 mm and the camera height is 718 mm: (a) the effect on the camera height changes on measuring relative height and target distance and (b) the effect on the relative height changes on measuring relative height and target distance.

The experimental results of camera I at the target distance of 3065 mm and the camera height of 718 mm are measured with a rangefinder. As shown in Table 3, multiple sets of experimental data can be obtained from five targets with the same spacing of 97 mm, and relative height and target distance can be calculated.

According to the multiple targets in an experimental image, the effects of different camera height changes on the measurement results can be obtained. The camera height changes are 197, 394, 591 and 788 mm respectively. The corresponding average relative height error and accuracy, and average target distance error and accuracy are shown in Figure 11(a). Where the minimum error of relative height is 6.47 mm, the minimum error of target distance is 63.01 mm, the maximum accuracy of relative height can reach 98.81%, and the maximum accuracy of target distance can reach 97.94%. The relative heights are 316, 513, 710, 907 and 1104 mm respectively. The corresponding average relative height error and accuracy, and average target distance error and accuracy are shown in Figure 11(b). Where the minimum error of relative height is 9.89 mm, the minimum error of target
When the target distance of camera I is 3065 mm, the camera height is changed to 656 mm, which further verifies the correctness of the experiment. As shown in Table 4, the experimental results of camera I at the target distance of 3065 mm and the camera height of 656 mm are measured with the rangefinder.

The same principle as above, the average relative height error and accuracy, and the average target distance error and accuracy corresponding to the camera height change can be shown in Figure 12(a). Where the minimum error of relative height is 0.77 mm, and the minimum error of target distance is 50.75 mm, the maximum accuracy of relative height can reach 99.88% and the maximum accuracy of target distance can reach 96.77%.

The average relative height error and accuracy, and the average target distance error and accuracy corresponding to the relative height can be shown in Figure 12(b). Where the minimum error of relative height is 21.95 mm, the minimum error of target distance is 54.46 mm, the maximum accuracy of relative height can reach 98.01% and the maximum accuracy of target distance can reach 96.65%.

The same experiment described above was carried out with Camera II. As shown in Table 5, the experimental results of Camera II at the target distance of 1123 mm and camera height of 718 mm are measured with the rangefinder.

The same principle as above, the average relative height error and accuracy, and the average target distance error and accuracy corresponding to the camera height change can be shown in Figure 13(a). Where the minimum error of relative height is 18.74 mm, the minimum error of target distance is 2.62 mm, the maximum accuracy of relative height can reach 97.28% and the maximum accuracy of target distance can reach 99.77%.

The average relative height error and accuracy, and the average target distance error and accuracy corresponding to the relative height can be shown in Figure 13(b). Where the minimum error of relative height is 14.54 mm,
the minimum error of target distance is 9.02 mm, the maximum accuracy of relative height can reach 97.02% and the maximum accuracy of target distance can reach 94.97%.

Camera II changes the camera height to 656 mm when the target distance is 1123 mm to further verify the correctness of the experiment. As shown in Table 6, the experimental results of Camera II at the target distance of 1123 mm and the camera height of 656 mm are measured with the rangefinder.

The same principle as above, the average relative height error and accuracy, and the average target distance error and accuracy corresponding to the camera height change can be shown in Figure 14(a). Where the minimum error of relative height is 15.02 mm, the minimum error of target distance is 1.83 mm, the maximum accuracy of relative height can reach 97.85% and the maximum accuracy of target distance can reach 99.84%.

The average relative height error and accuracy, and the average target distance error and accuracy corresponding to the relative height can be shown in Figure 14(b). Where the minimum error of relative height is 17.96 mm, the minimum error of target distance is 4.83 mm, the maximum accuracy of relative height can reach 96.83% and the maximum accuracy of target distance can reach 99.57%.

5.3. Error analysis

From Tables 3–6 and Figure 11–14, it can be seen that the measurement method is feasible and has high accuracy, which verifies the correctness of the theoretical analysis. From the imaging relationship Equation (19) for non-horizontal target by changing the camera height, it can be seen that the factors affecting the measurement results are the relative height of the measured target, the change of the camera height and the parameters of camera lens.

It can be seen from Table 7 that the error of camera II is generally smaller than that of camera I, because camera II is a zero distortion (Riley) camera. That is, camera distortion will affect the measurement accuracy. It can be seen from Table 8 that the stability of the measurement results of Camera II is worse than that of Camera I.
Figure 14. Diagram of experimental results of camera II when the target distance is 1123 mm and the camera height is 656 mm: (a) the effect on the camera height changes on measuring relative height and target distance and (b) the effect on the relative height changes on measuring relative height and target distance.

Table 4. Experimental results of camera I when the target distance is 3065 mm and the camera height is 656 mm.

| (h, m) | obj1        | obj2        | obj3        | obj4        | obj5        |
|-------|-------------|-------------|-------------|-------------|-------------|
| Original target | obj1 | (–,–) | (474.9, 3008.0) | (645.5, 3002.0) | (906.5, 3015.9) | (1094.9, 3014.2) |
|       | obj2 | (277.9, 3008.0) | (–,–) | (619.1, 3002.5) | (925.7, 3011.9) | (1104.8, 3010.9) |
|       | obj3 | (252.0, 3002.0) | (422.6, 3002.5) | (–,–) | (835.4, 2984.9) | (1149.1, 2995.1) |
|       | obj4 | (316.0, 3015.9) | (532.2, 3011.9) | (835.4, 2984.9) | (–,–) | (1071.2, 3021.9) |
|       | obj5 | (307.8, 3014.2) | (514.7, 3010.9) | (755.5, 2995.1) | (874.6, 3021.9) | (–,–) |

because Camera II is an ultra-wide-angle lens with camera focal length of 15 mm, which has wider field of view and is greatly affected by errors caused by segmentation. Due to the relatively simple experimental environment, the distance and height measured by the rangefinder will also have errors, which will affect the experimental accuracy to a certain extent.

5.4. Comparison with related work

The current research on visual measurement technology are mainly based on binocular or trinocular vision, or based on monocular vision with other auxiliary equipment, or based on monocular vision to measure horizontal target.

Ortiz et al. (2018) established a mathematical model for depth error by the ZED camera, which is a binocular vision system that derives corner points for known coordinates. Shao and Dong (2019) designed a trinocular vision sensor based on a camera with tilt-shift lens to enlarge the overlapping area for 3D reconstruction. Huang and Ye (2005) used a monocular camera and an auxiliary measurement probe to obtain 3D coordinates. Yuan and He (2020) used the distance from the person to the reference wall to derive the height of the person. Mao et al. (2020)
Table 5. Experimental results of camera II when the target distance is 1123 mm and the camera height is 718 mm.

| (h, m) | obj1 | obj2 | obj3 | obj4 | obj5 |
|-------|------|------|------|------|------|
| Original target | obj1 | (–, –) | (334.2, 1116.8) | (–, –) | (322.8, 1109.0) |
|          | obj2 | (–, –) | (500.8, 1108.1) | (–, –) | (549.5, 1121.1) |
|          | obj3 | (–, –) | (763.1, 1103.2) | (–, –) | (710.0, 1123.0) |
|          | obj4 | (–, –) | (–, –) | (–, –) | (–, –) |
| Truth value | (316.0, 1123.0) | (513.0, 1123.0) | (710.0, 1123.0) | (907.0, 1123.0) | (1104.0, 1123.0) |

Table 6. Experimental results of camera II when the target distance is 1123 mm and the camera height is 656 mm.

| (h, m) | obj1 | obj2 | obj3 | obj4 | obj5 |
|-------|------|------|------|------|------|
| Original target | obj1 | (–, –) | (349.0, 1122.3) | (–, –) | (324.2, 1109.5) |
|          | obj2 | (–, –) | (619.0, 1114.1) | (–, –) | (494.5, 1114.1) |
|          | obj3 | (–, –) | (890.9, 1113.6) | (–, –) | (496.0, 1114.1) |
|          | obj4 | (–, –) | (993.3, 1043.6) | (–, –) | (543.0, 1121.9) |
| Truth value | (316.0, 1123.0) | (513.0, 1123.0) | (710.0, 1123.0) | (907.0, 1123.0) | (1104.0, 1123.0) |

Table 7. Error analysis of experimental results of camera I and camera II.

| Camera I | Camera II |
|----------|-----------|
| Unit: mm | Relative height | Target distance | Relative height | Target distance |
| Average error | 31.35 | 62.28 | 28.60 | 17.37 |
| Maximum error | 125.44 | 80.11 | 100.99 | 79.36 |
| Minimum error | 0.03 | 43.07 | 0.02 | 0.71 |

Table 8. Stability analysis of experimental results of camera I and camera II.

| Standard deviation | Camera I (m = 3065 mm) | Camera II (m = 1123 mm) |
|--------------------|------------------------|------------------------|
| Relative height | Target distance | Relative height | Target distance |
| h_r = 718 mm | 19.58 | 4.00 | 29.11 | 21.45 |
| h_r = 656 mm | 39.71 | 10.61 | 23.85 | 22.65 |

6. Conclusion

Based on the imaging relationship between height and distance of non-horizontal target, a geometric model of the imaging relationship is derived, and a new monocular vision measurement method for non-horizontal target is proposed. According to the relevant external parameters (camera height) and internal parameters (focal length, photosensitive element size, image size), and target self-height, the equations for calculating relative height and target distance are obtained. The experimental results show that the accuracy of relative height measurement and accuracy of target distance measurement based on monocular vision can reach 95.88% and 98.45% on average. The proposed method is simple and effective. Without data collection and complicated training like machine learning, the required target distance and target height can be quickly calculated. For further research, we can focus on the measurement study of partially obscured target, judge the integrity and consistency of the target and then measure it, or directly measure part of the target.

Table 9. Table of comparison with related methods.

| Methods | Number of cameras | Other auxiliary equipment | Measure horizontal target | Measure non-horizontal target | Measure target distance |
|---------|------------------|--------------------------|---------------------------|-------------------------------|------------------------|
| Ortiz et al. (2018) | 2 | ✓ | ✓ | × | ✓ |
| Shao and Dong (2019) | 3 | × | ✓ | ✓ | ✓ |
| Huang and Ye (2005) | 1 | ✓ | ✓ | × | ✓ |
| Yuan and He (2020) | 1 | ✓ | ✓ | × | × |
| Mao et al. (2020) | 1 | × | ✓ | × | ✓ |
| Ours | 1 | × | ✓ | ✓ | ✓ |

realized the depth analysis of the horizontal target of the monocular camera according to the geometric model of the camera imaging and the imaging parameters of the camera.

Table 9 lists the methods adopted by the above technologies. Compared with their method, our method has three advantages: (1) simple algorithm, few devices, low computational complexity and easy to implement; (2) wide range of applications, which can measure horizontal and non-horizontal targets; (3) accurate calculation of relative height and target distance.
There is also a direct effect of the camera’s rotation on the measurement results, which can also be further studied.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by Zhejiang province Natural Science Foundation of China: [Grant Number LY20F020022]; National Natural Science Foundation of China: [Grant Number 62176237].

**References**

Chen, D. K. (2021). Active disturbance rejection control of indoor inspection robot in intelligent substation based on monocular vision. *Journal of Physics: Conference Series*, 1894(1), 012042. https://doi.org/10.1088/1742-6596/1894/1/012042

Cheng, X., Sun, J., Zhou, F., Xiaoqi, C., Junhua, S., Fuqiang, Z. (2021). Shape from apparent contours for bent pipes with constant diameter under perspective projection. *Measurement*, 182, 109787. https://doi.org/10.1016/j.measurement.2021.109787

Ferro, M., Paolillo, A., Cherubini, A., et al. (2019). Vision-based navigation of omnidirectional mobile robots. *IEEE Robotics and Automation Letters*, 4(3), 2691–2698. https://doi.org/10.1109/LRA.2019.2913077

Ge, P., Wang, H., Zhang, Q., et al. (2021). Mirror-assisted multi-view DIC for 360° panoramic large deformation measurement. *Optics and Lasers in Engineering*, 146, 106673. https://doi.org/10.1016/j.optlaseng.2021.106673

Haertel, M. E. M., Pinto, T. d. C., & Júnior, A. A. G. (2015). Trinocular stereo system with object space oriented correlation for inner pipe inspection. *Measurement*, 73, 162–170. https://doi.org/10.1016/j.measurement.2015.05.015

Hu, O. J., Feng, Z. Y., He, L. P., et al. (2020). Accuracy improvement of binocular vision measurement system for slope deformation monitoring. *Sensors*, 20(7), 1994. https://doi.org/10.3390/s20071994

Huang, G. P., & Ye, S. H. (2005). Principle of coordinates acquisition based on single camera. *Transactions of Tianjin University*, 11(1), 49–52.

Javanmardi, M., & Qi, X. J. (2018). Visual tracking of resident space objects via an RFS-based multi-Bernoulli track-before-detect method. *Machine Vision and Applications*, 29(7), 1191–1208. https://doi.org/10.1007/s00138-018-0963-6

Jiehu, K., Bin, W., & Yuchan, W. (2020). Calibration method of 1D laser rangefinder for large-scale 3D measurement under harsh environment. *IEICE Electronics Express*, 17(8), 1–5. https://doi.org/10.1587/elex.17.20200071

Kong, S. H., Fang, X., Chen, X. Y., et al. (2020). A NSGA-II-based calibration algorithm for underwater binocular vision measurement system. *IEEE Transactions on Instrumentation and Measurement*, 69(3), 794–803. https://doi.org/10.1109/TIM.2019.2906974

Liu, L. (2021). Lane offset survey for one-lane horizontal curves using binocular stereo vision measurement system. *Journal of Surveying Engineering*, 147(4), 04021017. https://doi.org/10.1061/(ASCE)SU.1943-5428.0000370

Liu, X. C., Yan, Y., & Gan, H. Y. (2021). Research on driverless vehicle vision algorithm// Proceedings of the MATEC Web of conferences. Hainan, China, 06001.

Mansour, M., Davidson, P., Stepanov, O., et al. (2019). Relative importance of binocular disparity and motion parallax for depth estimation: A computer vision approach. *Remote Sensing*, 11(17), 1990. https://doi.org/10.3390/rs11171990

Mao, J. F., Huang, W., & Sheng, W. G. (2020). Target distance measurement method using monocular vision. *IET Image Processing*, 14(13), 3181–3187. https://doi.org/10.1049/iet-ipr.2019.1293

Mao, J. F., Zhang, M. G., Zhu, L. N., et al. (2017). Target depth measurement for machine monocular vision//Proceedings of the Pacific-Rim conference on multimedia (PCM), Harbin, China, 583–595.

Martínez-Martin, E., & Del Pobil, A. P. (2021). Robot vision for manipulation: A trip to real-world applications. *IEEE Access*, 9, 3471–3481. https://doi.org/10.1109/ACCESS.2020.3048053

Martínez-Diaz, S. (2021). 3D distance measurement from a camera to a mobile vehicle, using monocular vision. *Journal of Sensors*, 2021, 1–8. https://doi.org/10.1155/2021/5526931

Ortiz, L. E., Cabrera, E. V., & Goncalves, L. (2018). Depth data error modeling of the ZED 3D vision sensor from stereobas. *Electronic Letters on Computer Vision and Image Analysis*, 17(1), 1–15. https://doi.org/10.1056/rev/elecvia.1084

Ou, J. J., Guo, X., Zhu, M., et al. (2021). Autonomous quadrotor obstacle avoidance based on dueling double deep recurrent Q-learning with monocular vision. *Neurocomputing*, 441, 300–310. https://doi.org/10.1016/j.neucom.2021.02.017

Rani, M. R., Mustafar, M. Z. C., Ismail, N. H. F., et al. (2021). Road peculiarities detection using deep learning for vehicle vision system. *IOP Conference Series: Materials Science and Engineering*, 168(1), 012001. https://doi.org/10.1088/1757-899X/1688/1/012001

Reina, G., Underwood, J., Brooker, G., et al. (2011). Radar-based perception for autonomous outdoor vehicles. *Journal of Field Robotics*, 28(6), 894–913. https://doi.org/10.1002/rob.20393

Riley, J. Venus Laowa 15mm f/2 FE Zero-D review. https://ephotozine.com/article/venus-laowa-15mm-f-2-fe-zero-d-review-31801, 2018-1-1

Shankar, A., Vatsa, M., & Sujit, P. B. (2018). A low-cost monocular vision-based obstacle avoidance using SVM and optical flow. *Unmanned Systems*, 6(4), 267–275. https://doi.org/10.1142/S2301385018500097

Shao, M. W., & Dong, J. Y. (2019). Trinocular vision system based on cameras with tilt–shift lens. *Journal of Optics*, 48(1), 43–48. https://doi.org/10.1088/1557-415X/12/5/012024

Song, Y. Y., Sheng, Y. Y., Chao, L., et al. (2021). Autonomous obstacle avoidance scheme using monocular vision applied to mobile robots. *Journal of Physics: Conference Series*, 1894(1), 012024. https://doi.org/10.1088/1742-6596/1894/1/012024

Yao, Z. B., Xie, J. Q., Tian, Y. Q., et al. (2019). Using Hampel identifier to eliminate profile-isolated outliers in laser vision measurement. *Journal of Sensors*, 2019, 1–12. https://doi.org/10.1155/2019/3823691

Yuan, F., & He, J. (2020). Human height measurement in surveillance video based on vision technology. *International Core Journal of Video Engineering*, 6(5), 198–208. doi:10.6919/ICJE.2020.005_605.0030