UDC 004.021

METHOD FOR DUAL-AGENT COLLABORATIVE MONOCULAR RANGING

Q.C. GUO, M.I. ZORKO, E.G. MAKEICHIK, V.Yu. TSVIATKOU

Belarusian State University of Informatics and Radioelectronics, Republic of Belarus

Received March 31, 2026

Abstract. In unknown environments, a single monocular camera struggles to achieve absolute distance measurement due to the lack of scale information. This paper proposes a vision-based ranging system utilizing two mobile agents that collaborate through mutual calibration, coordinated observation, and triangulation to estimate the absolute distance to an arbitrary target. The system operates in three stages: cooperative initialization and baseline calibration, angular disparity measurement, and triangulation-based distance calculation. We specifically introduce a chain tracking measurement method to extend the effective field of view, addressing the challenge when the target and reference point are not simultaneously observable. Experimental results demonstrate that the proposed method achieves acceptable ranging errors while maintaining a favorable balance between accuracy and applicability.

Keywords: Monocular vision, collaborative ranging, triangulation, dual-agent collaboration, multi-agent system, unknown scale.

Introduction

With the widespread deployment of mobile robots and unmanned aerial vehicles in applications such as inspection, search and rescue, and warehousing, environmental perception and target localization have become critical challenges [1]. Monocular cameras are widely adopted due to their small size, low cost, and rich information output [2]; however, their inherent scale ambiguity limits their ability to perform absolute distance measurement.

Existing approaches often rely on prior maps, known target dimensions, or multi-sensor fusion (e.g., LiDAR, IMU), which face limitations in resource-constrained or unknown environments. Although binocular stereo vision can recover scale, it requires precise calibration and a fixed baseline, offering limited flexibility [3].

This paper presents a collaborative visual ranging method based on two mobile agents, each equipped with a monocular camera. The agents determine the baseline through mutual calibration [4] and then apply the law of sines using pixel positions of the target and reference points to compute the target distance. We also design a chain tracking measurement strategy [5] to overcome the limitation of the physical field of view, making the method suitable for wide-baseline scenarios. The approach is presented using a dual-agent configuration as the fundamental unit but can be naturally extended to multi-agent systems: any two agents with communication and collaboration capabilities can apply the same strategy for target ranging.

Problem Description and Model Formulation

This work aims to address the challenge that a single monocular camera cannot determine distance when the target's scale is unknown in an environment without prior information. We propose a collaborative visual ranging method using two mobile agents. The system consists of two mobile agents (Vehicle A and Vehicle B), each equipped with a monocular camera. Through mutual calibration, coordinated observation, and triangulation, the system estimates the absolute distance to an arbitrary target C.

The fundamental assumptions are as follows:

1. Both cameras A and B have been intrinsically calibrated. The camera equivalent focal length f, principal point (u_0, v_0) distortion coefficients, and horizontal field of view FOV_h are known.

2. Each camera is mounted on a square baseplate with a known side length $L \times L$. The vertical offset between the camera's optical center and the baseplate center is negligible or has been compensated for, meaning the baseplate center is considered coincident with the camera's optical center.

3. Stable communication exists between the two agents, allowing the exchange of control commands and image feature data.

4. The environment is static or quasi-static; the relative positions of A, B, and C remain unchanged during a single distance measurement process.

Methodology

The workflow of the proposed collaborative ranging system is divided into three stages: Collaborative Initialization and Baseline Calibration, Viewing Angle Difference Measurement, and Triangulation-based Distance Calculation.

1. Stage one: collaborative initialization and baseline distance calibration.

When agent A detects the target object C and sends an assistance request to B, both agents rotate in place via a control strategy until they capture each other's square baseplates in their images.

1.1. Establishing mutual gaze: Image processing algorithms (such as edge detection, Hough transform, or deep learning-based object detection) identify the imaging contour of the partner's baseplate. Control algorithms adjust the pan-tilt unit or vehicle body attitude until the image center of the partner's baseplate aligns with the agent's own image center. At this point, the optical axes of the two cameras are considered approximately coincident along the same central axis, establishing a "mutual gaze" state.

1.2. Baseline distance calculation: In the mutual gaze state, agent A (or B) calculates the imaging size of the partner's baseplate (agent B or A) through its image. Let the real side length of the square baseplate be L, and the detected pixel width of the border in the image be w(pixels). As shown in Figure 1, according to the camera's pinhole model and known intrinsic parameters, the distance from agent B's baseplate to camera A, i.e., the baseline distance AB, can be calculated as

$$AB = \frac{f \cdot L}{w}, \tag{1}$$

where f is the equivalent focal length after camera calibration.

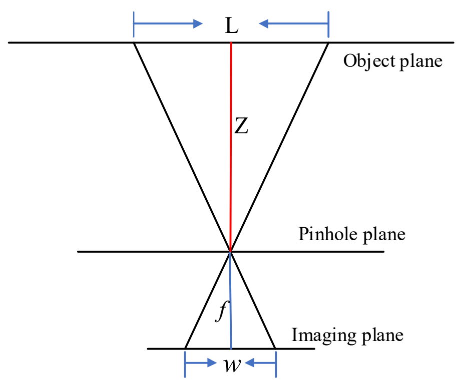


Figure 1. Pinhole Camera Model

2. Stage two: Viewing angle difference measurement.

This stage aims to accurately measure $\angle BAC$ and $\angle ABC$. The two agents take turns acting as the observer, using symmetrical methods. The measurement of $\angle ABC$ is detailed below using agent B observing A and C as an example.

2.1. Pixel-to-angle conversion model:

Given the camera's horizontal resolution W (pixels) and horizontal field of view FOV_h, the horizontal angular resolution per pixel θ_p is

$$\theta_p = \frac{FOV_h}{W}, \tag{2}$$

this model is the foundation for subsequent angle calculations.

2.2. Viewing angle difference measurement strategy:

Agent B starts rotating in place (e.g., clockwise) from its mutual gaze posture with A to search for target C. Depending on whether C and A can be simultaneously within B's field of view, two cases are considered:

Case 1: Co-visibility measurement.

If during rotation, both A and C appear simultaneously in B's view (as shown in Figure 2a), directly extract the horizontal pixel coordinates u_A and u_C of the centroids of A and C in the image. The viewing angle difference $\angle ABC$ is

$$\angle ABC = |u_C - u_A| \cdot \theta_p . \tag{3}$$

Case 2: Feature point chain tracking measurement.

Typically, because the baseline AB is relatively short compared to distance BC, the angular separation between points A and C in agent B's view may be very large, exceeding the camera's FOV, preventing simultaneous observation. To address this, we propose a chain tracking measurement scheme (as shown in Figure 2b). The specific process is as follows:

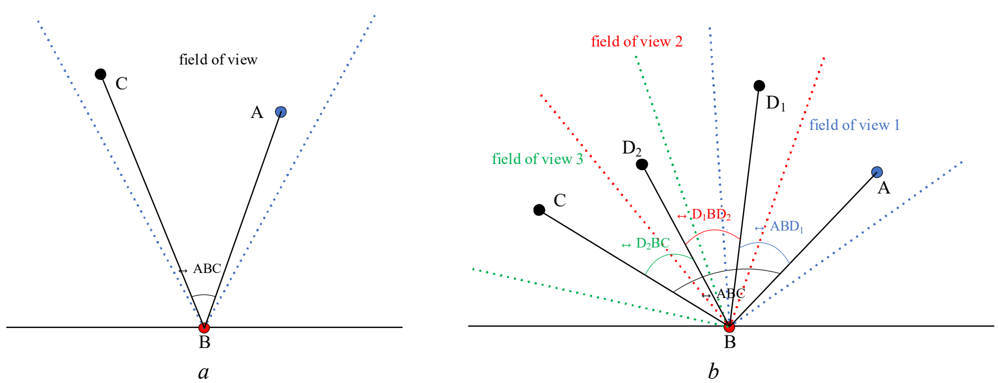


Figure 2. Visible situation: a – Simultaneously visible;
b – Feature Point Chain Tracking Measurement

(1) Agent B begins rotating. When A moves to the left edge of its view (a preset threshold) and C has not yet appeared, it searches for a stable, easily trackable natural feature point D_1 near the right edge of the view.

(2) Record the horizontal pixel coordinates u_A and u_{D_1} of A and D_1 at this moment, calculate and store the angle $\angle ABD_1$ is

$$\angle ABD_1 = |u_{D_1} - u_A| \cdot \theta_p . \tag{4}$$

(3) Agent B continues rotating, using D_1 as the new reference. When D_1 moves to the left edge of the view and C still hasn't appeared, similarly find the next strong feature point D_2 near the right edge.

(4) Record the horizontal pixel coordinates u_{D_1} and u_{D_2} of D_1 and D_2 at this moment, calculate and store $\angle D_1BD_2$ is

$$\angle D_1BD_2 = \left|u_{D_1} - u_{D_2}\right| \cdot \theta_p . \quad (5)$$

(5) Repeat this process until the target C appears within the view. Record the pixel coordinates of the last feature point D_i and C, calculate $\angle D_iBC$ is

$$\angle D_iBC = \left|u_C - u_{D_i}\right| \cdot \theta_p . \quad (6)$$

(6) Finally, $\angle ABC$ is the algebraic sum of all intermediate angles

$$\angle ABC = \angle ABD_1 + \angle D_1BD_2 + \cdots + D_iBC . \quad (7)$$

This method effectively extends the virtual field of view by introducing a series of intermediate natural feature points as "angular beacons," overcoming the limitation of the physical field of view.

Similarly, agent A employs the same strategy to rotate and measure $\angle BAC$.

3. Stage three: triangulation-based distance calculation

After obtaining the baseline distance AB and the two angles $\angle BAC$ and $\angle ABC$, all parameters of triangle ABC are determined. According to the trigonometric sine law, the lengths of the other two sides, i.e., the distances from target C to the two agents, can be calculated as

$$AC = \frac{AB \cdot \sin(\angle ABC)}{\sin(180^\circ - \angle BAC - \angle ABC)} = \frac{AB \cdot \sin(\angle ABC)}{\sin(\angle BAC + \angle ABC)} . \quad (8)$$

Thus, we accomplish absolute distance measurement to the target object using only two collaborative monocular cameras, without prior knowledge of the object's size.

Experiments and Analysis

To validate the effectiveness of the proposed collaborative ranging method, we analyze the theoretical error of the target distance AC based on first-order error propagation theory. The comparative experiments are conducted using the baseline system parameters listed in Table 1.

Table 1. **System basic parameters**

Parameter (Symbol)	Value (Unit)
Agent size (*L*)	0.05(m)
Camera equivalent focal length (*f*)	3063 (pixels)
Image horizontal resolution (*W*)	4032 (pixels)
Camera horizontal field of view (FOV_h)	70 (deg)

1. Theoretical error model.

The ranging error of the proposed method originates mainly from baseline calibration and angular measurement, which propagate through the triangulation calculation. According to first-order error propagation theory, the maximum relative error of AC is given by

$$\delta AC_{\max} = \frac{AB}{f \cdot L} + \left|\cot(\angle BAC + \angle ABC)\right| \cdot k\theta_p + \frac{\sin(\angle BAC)}{\sin(\angle ABC)\sin(\angle BAC + \angle ABC)} \cdot k'\theta_p , \quad (9)$$

where k and k' are the chain tracking steps, and θ_p is the angular resolution per pixel (in degrees/pixel). This expression captures the combined effects of baseline length, label size, camera parameters, and geometric configuration on ranging accuracy.

For a calibrated standard binocular stereo vision system with parallel optical axes, baseline length B, and equivalent focal length f, we define a coordinate system with the left camera's optical center A as the origin: the X-axis is along the baseline direction (positive to the right), the Z-axis is along the optical axis (positive forward), and the Y-axis is perpendicular to the plane (downward). For simplicity, the target is assumed to lie in the XZ plane (Y=0). The coordinates of point C are (X, Z). Using first-order error propagation, the maximum relative error of the slant distance AC is

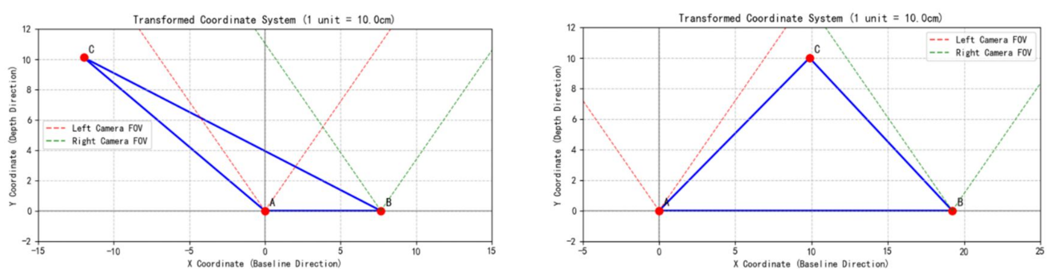
$$\delta AC_{\max\_stereo} = \frac{|X| \cdot Z}{2f(X^2 + Z^2)} + \frac{Z}{f \cdot B}. \quad (1)$$

2. Experimental comparison and analysis.

To assess the performance of the proposed method (dual-agent collaborative monocular vision) and binocular stereo vision in unknown environments, we conducted large-scale Monte Carlo simulations. The experiment comprised 100,00 independent trials, each generating 20 random triangles satisfying geometric constraints (baseline 50 - 500 cm, target distance 100 - 1000 cm, angles 10° - 170°), resulting in a total of 200,000 scenarios. The measurability of the binocular system and the theoretical relative errors of both methods were statistically analyzed.

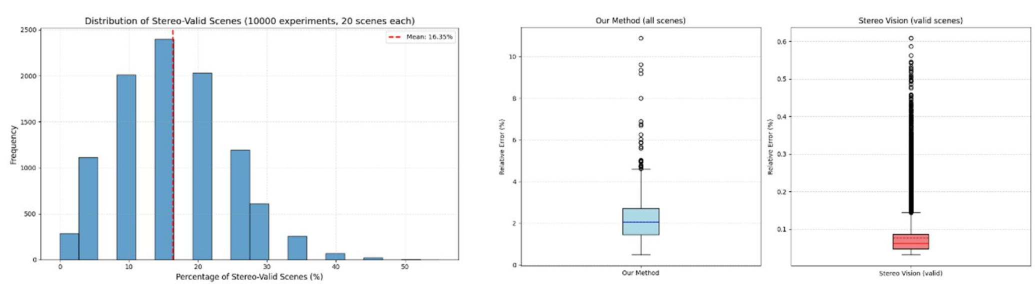
Regarding applicability, the proposed method demonstrates a significant advantage over traditional binocular stereo vision. As illustrated in Fig. 3, there are two types of scenarios where the binocular system fails because the target lies outside the common field of view.

Statistical results show that the average proportion of measurable scenes for the binocular system is only 16.42%, with a standard deviation of 8.36%. In the worst case, none of the 20 scenes in a trial were measurable (0% measurability), while the best case achieved only 55.0% (Fig. 4a). This indicates that in typical unknown environments, the applicability of binocular stereo vision is severely limited by the common field-of-view constraint, with over 80% of random geometric configurations being unsuitable for binocular ranging. In contrast, the proposed method achieves 100% scene measurability through its chain tracking strategy.



a *b*

Figure 3. Situations that cannot be measured under a binocular system



a *b*

Figure 4. Experimental results: a – the proportion that can be measured by the binocular system; b – error rate

In terms of accuracy, the proposed method successfully estimated distances in all 200,000 scenarios without failure. Figure 4b shows that the error distribution of the proposed method is centered around 2%, with the median close to the mean, indicating a symmetric distribution. The error range spans 0.48% to 10.89%, covering typical scenarios from close range to long distance and various angles. This accuracy meets the basic requirements of many robotic tasks (e.g., obstacle avoidance, target approach). For scenes where binocular vision is applicable, its error is extremely small (mean 0.076%) with a highly concentrated distribution (median 0.062%), demonstrating high and stable accuracy under favorable geometric conditions. However, its main limitation remains the restricted set of measurable scenes, while the proposed method achieves a favorable balance between accuracy and applicability.

Conclusion

This paper presents a monocular vision ranging system based on dual-agent collaboration, capable of absolute distance estimation for arbitrary targets in unknown environments. The system operates through three stages: cooperative initialization and baseline calibration, angular disparity measurement, and triangulation-based distance calculation, enabling ranging without prior knowledge of the target's size. Notably, the proposed chain tracking measurement strategy effectively extends the virtual field of view, addressing the challenge when the target and reference point cannot be observed simultaneously, thus making the method suitable for wide-baseline applications.

The method is described using a dual-agent configuration as the minimal implementation unit, but its core mechanisms (mutual calibration, collaborative observation, and triangulation) are fully applicable to multi-agent systems. In such systems, any two agents with communication and collaboration capabilities can apply the same strategy for collaborative ranging, thereby forming a distributed collaborative perception network and enhancing perception coverage in large-scale unknown environments.

Future work will focus on real-time collaborative ranging in dynamic scenes, multi-agent collaborative optimization, and the deployment and validation of the method on practical robotic platforms.

References

1. Lyu M., Zhao Y., Huang C., et al. Unmanned aerial vehicles for search and rescue: A survey. // Remote Sensing. 2023. Vol. 15. P. 3266.
2. Guo Q. Implementation principle of monocular ranging. // Технологии передачи и обработки информации : материалы Международного научно-технического семинара. Минск, 2024. P. 78–82.
3. Guo Q. Research on relative localization of mobile robots based on monocular vision and known tags. // Технологии передачи и обработки информации : материалы Международного научно-технического семинара. Минск, 2025. P. 141–143.
4. Guo Q. Research on depth estimation for monocular cameras based on prior information. // Технологии передачи и обработки информации : материалы Международного научно-технического семинара. Минск, 2025. P. 179–182.
5. Guo Q. Method for extended field-of-view angle measurement based on chain tracking of feature points. // XII Белорусско-Китайский молодежный инновационный форум «Новые горизонты - 2025». Минск, 2025. P. 72–74.