Practical Method of Low-Light-Level Binocular Ranging Based on Triangulation and Error Correction

Qi Shi
Institution of Automation, Chinese Academy of Sciences
No. 95, Zhongguancun East St.
Beijing, China
+86 13261537853
shiqi2014@ia.ac.cn

Lei Ma
Institution of Automation, Chinese Academy of Sciences
No. 95, Zhongguancun East St.
Beijing, China
lei.ma@ia.ac.cn

Yiping Yang
Institution of Automation, Chinese Academy of Sciences
No. 95, Zhongguancun East St.
Beijing, China
yiping.yang@ia.ac.cn

ABSTRACT
Reliable target distance measurement at long range during the night is crucial for security surveillance at night. We present a practical method for low-light-level ranging using a motionless stereo vision system. The approach can measure the target distance within 200 meters. Our range finding algorithm uses triangulation to reconstruct target point coordinates, instead of precomputing disparity maps. Trilateration method is adopted to calculate median distance of the target. Based on statistical tests, a biexponential-function-based error correction model is proposed. The distance estimation is corrected to final distance after fitting the model. Ranging error can be dramatically reduced by error correction. Extensive experiments on datasets show that our method achieves excellent ranging performance and is stable in repeatability, the mean relative accuracy is below 1%.

CCS Concepts
• **General and reference → Measurement** • **Computing methodologies → Reconstruction**

Keywords
Low-Light-Level Image; Binocular Stereo Vision; Triangulation; Trilateration; Error Correction

1. INTRODUCTION
With the development of optoelectronic technique, various kinds of detection and measurement methods have been put into use. Measuring the target distance is one of the most important task amongst them. According to the special need of concealed ranging at night, to measure distances passively in low-light-level circumstance is of great importance. Passive ranging uses natural light radiation of the target to measure distance. At present, binocular ranging is most widely used, which usually retrieves the depth map by matching two images of FOV (Field of View) collected by two cameras.

According to the principle of binocular parallax, we usually use calibration, rectification and stereo matching to get stereo images that are horizontally aligned, the depth is derived from disparity map[1]. The depth is negatively correlated to disparity, that is, when an object is far away from stereo cameras, the error will increase significantly. But in our application, the measurable range needs to be 200 meters in low-light-level environment, which will cause large error and can't be applied in practice.

3-d reconstruction is another way to get depth. We solve the essential matrix through a bunch of matching points in stereo images, then triangulate the space coordinates of the matching points. But the algorithm is sensitive to noise, which is inevitable in low-light-level circumstance. Besides, in 200-meter distance, the accuracy of reconstruction will be seriously reduced. Thus, reconstruction method is not capable of solving this problem.

Besides binocular methods mentioned above, there are some monocular approaches. A few methods have been developed, such as structured light method[2], Aided Measuring Probe (AMP) method, geometric shape constraint method[3], and so on. However, they rely on the high resolution along with great focusing and zooming performance of the camera, which are not provided by low-light-level ones, limiting the capability in solving the problem of low-light-level ranging.

In error correction of distance measurement, common methods can be classified into two kinds. One is correcting the parameters of vision system, such as improving calibration method for specific scene. The other is modeling the influence of camera position on distance value, such as introducing rotation and translation of cameras to the distance measurement formula[4] or conduct stereo rectification.

To address the shortcomings of previous methods and to meet the need of 200-meter ranging in low-light-level circumstance, we consider applying triangulation and error correction to improve the accuracy. In our method, ranging error can be stably estimated, and accurate measurement can be obtained by error correction. To our knowledge, it is the first time this pipeline is used to solve the problem.

This paper is structured as follows: In Section 2, we introduce a practical new method of low-light-level binocular ranging. In Section 3, experiment results are displayed, including introduction to the experimental environment. Finally, conclusions are made in Section 4.

2. OUR APPROACH
In the low-light-level circumstance, images captured by camera have lower quality as well as severe noise, makes traditional methods that require high image quality inapplicable.

Pipeline of our method is shown in Figure 1. The first step is to calibrate the cameras and capture training images in a stationary camera position. Next, we conduct feature extraction and feature matching on ROI to get matching points on target. Matching

CSAI 2017, December 5–7, 2017, Jakarta, Indonesia

ACM ISBN 978-1-4503-5392-2/17/12…$15.00
DOI: https://doi.org/10.1145/3168390.3168417

points are triangulated, followed by trilateration to get ranging estimation. Finally, the error correction model is fitted using the ranging results acquired. The corrected result forms the final results.

2.1 Notations

Binocular vision system is composed of two cameras with same focal length and working status, as shown in Figure 2. O_l and O_r are the optical centers of the two cameras. The origin of space coordinates system coincides with O_l. Denote target point to be measured as P, p_l and p_r are the corresponding image points of P.

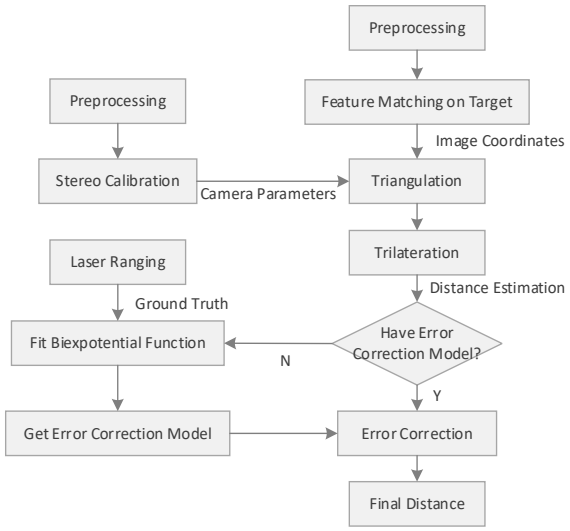


Figure 1. Pipeline of our approach.

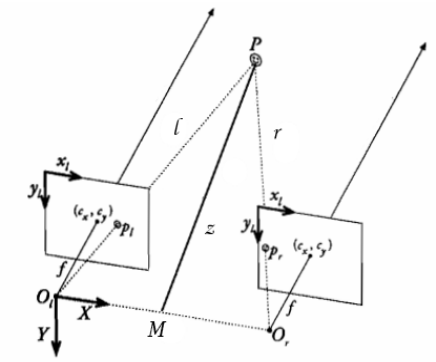


Figure 2. Stereo vision system architecture.

2.2 Preprocessing

Due to the low signal-to-noise ratio (SNR) and low contrast of low-light-level images, preprocessing is needed.

In low-light-level images, stripe noise and impulse noise are the major noise types. We conducted bilateral filtering[5] to reduce them, because it can keep the edge information while denoising, which is beneficial to calibration in low-light-level circumstances and preserve more target features while acquiring ranging images.

Besides noise suppression, image enhancement is also required. We adopt contrast limited adaptive histogram equalization (CLAHE) to enhance the images[6]. To limit the overamplification of local noises, CLAHE use predefined threshold to clip the histogram before computing cumulative distribution function (CDF), which is suitable of low-light-level images. As is shown in Figure 3, the image quality has been enhanced significantly after preprocessing.

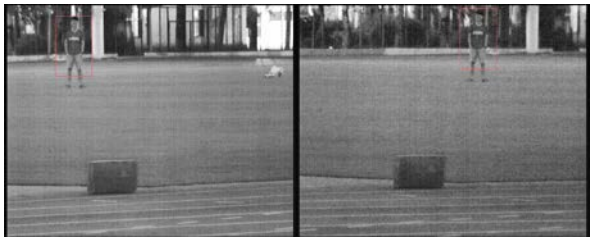
2.3 Calibration and Feature Matching

The intrinsic and extrinsic parameters of cameras should be calibrated first. The flexible calibration method proposed in [7] is used. A planar calibration board is used for the method (See Figure 3), the cameras observe the planar pattern from some different orientations. Focal length f, principle point C and distortion model are obtained by the calibration algorithm. Rotation matrix and translation vector are obtained by stereo calibration.

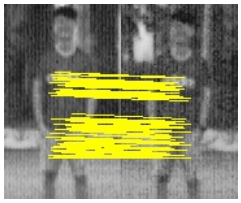
(a) (b)

Figure 3. Enhancement and noise reduction of low-light-level images: (a) Input low-light-level image. (b) Image after preprocessing.

To get accurate match points of ranging target, we use ORB[8] feature and Grid-based Motion Statistics (GMS) [9] to select corresponding features. In 200-meter distance, targets are usually small, so we only conduct feature matching in an ROI. Match points and ROIs in images are shown in Figure 4.



(a) (b)



(c)

Figure 4. Feature matching of low-light-level images in ROIs: (a) Input low-light-level image (left camera). (b) Input low-light-level image (right camera). (c) GMS feature matching results, corresponding points are connected by a line.

After feature matching, we have obtained a set of matching points of our ranging target. Next, triangulation and trilateration are performed to estimate target distance.

2.4 Triangulation and Trilateration

By analyzing the corresponding relations of image points, we can get spatial coordinates of the target, which are used to measure target distance in physical space.

Triangulation[10] is to find the space coordinates of a point given its position in two images taken by cameras with known calibrations and poses. After stereo calibration and feature matching, camera matrix and pose along with target position in the two images become known, then simple linear triangulation method is used to get the position of target in spatial coordinates.

Suppose the homographic matrix of one camera $H = K[R \quad t]$, where K is the camera matrix, and R, t are rotation matrix and translation vectors between left and right camera coordinate systems. Given spatial point P on target, denote image point of P as p, there is the projection relation:

$$p = HP \tag{1}$$

In equation (1), $p = s(u, v, 1)^T$, where s is a scale parameter, and $P = (X, Y, Z, 1)^T$. Denoting H_i^T the i th row of the matrix H, the equation may be written as:

$$\begin{cases} su = H_1^T P \\ sv = H_2^T P \\ s = H_3^T P \end{cases} \tag{2}$$

Eliminating s using the third equation, we arrive at:

$$\begin{cases} \left(uH_3^T - H_1^T\right)P = 0 \\ \left(uH_3^T - H_2^T\right)P = 0 \end{cases} \tag{3}$$

For left and right camera respectively, we obtain a total of four linear equations through equation (2). Use inhomogeneous solution, write P in inhomogeneous coordinates and rearrange the equation in the form of $AP = B$:

$$\begin{bmatrix} u_l K_{l3}^T - K_{l1}^T \\ v_l K_{l3}^T - K_{l2}^T \\ u_r (K_r R_r)_3^T - (K_r R_r)_1^T \\ v_r (K_r R_r)_3^T - (K_r R_r)_2^T \end{bmatrix} P = - \begin{bmatrix} 0 \\ 0 \\ u_r (K_r t_r)_3^T - (K_r t_r)_1^T \\ u_r (K_r t_r)_3^T - (K_r t_r)_2^T \end{bmatrix} \tag{4}$$

In equation (4), the index l and r denotes image points in left and right camera. A is 4×3 matrix, B is 4×1 vector. These equations define P only up to an indeterminant scale factor, and we seek a nonzero solution for P:

$$P = \left(A^T A\right)^{-1} A^T B \tag{5}$$

With noisy data, the equations will not be satisfied precisely, and we seek a best solution through least-squares solutions.

The geometric meaning of triangulation is to find an intersection point of extension cord of $O_l p_l$ and $O_r p_r$ in space coordinates. When the space coordinates of target point are determined, the distance of the target can be measured.

Typically, we use vertical distance to do the ranging, but in 200-meter circumstance, vertical distance to the image plane is not capable to be the evaluation index, because the target has a large moving range in the FOV. As shown in Figure 5, P and Q have the same vertical distance ($h_1 = h_2$), but their straight-line distance to the stereo system are different ($h_3 > h_2$) and this error cannot be neglected. Besides, when using vertical distance, ground truth may not be obtained easily to evaluate the algorithm.

Because of this, we use median distance z to measure the object distance, which is the length of target point P and midpoint M of the baseline (See Figure 2). Trilateration is adopted to calculate z:

$$z = \frac{1}{2}\sqrt{2l^2 + 2r^2 - d^2} \tag{6}$$

Equation (6) is derived from law of cosines of $\Delta O_l PM$ and $\Delta O_l P O_r$. l and r are the distance between P and O_l, O_r, d is baseline distance.

As the origin of spatial coordinates system coincides with O_l, l can be obtained by knowing the space coordinates of P:

$$l = \|P\|_2 \tag{7}$$

r is calculated through the right camera coordinates of point P. Due to the pose of stereo cameras, we get:

$$\begin{cases} P_r = RP_l + t = RP + t \\ r = \|P_r\|_2 \end{cases} \tag{8}$$

In equation (8), P_l and P_r are left and right camera coordinates of P. d can be obtained directly by translation vector between the coordinate systems of the cameras:

$$d = \|t\|_2 \tag{9}$$

l, r, d are calculated from the results of calibration and triangulation. Substitute into equation (6), median distance is obtained.

In Chapter 2.3, we get a set of matching points through feature matching. When conducting triangulation and trilateration, matching points obtained from the background can result in false distance. In practice, the abrupt difference in foreground-background disparities can be effectively suppressed by choosing the median of all ranging results.

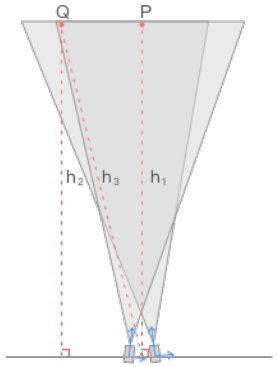


Figure 5. The target has a large moving range in the FOV.

2.5 Error Correction

Up till now, a rough estimation of target distance is gotten, to get the distance refinement model, we have conducted 7 experiments at night with different camera positions, as shown in Figure 6.

Under low-light-level circumstances, the imaging quality of industrial CCD camera is highly confined by noise level, so low-light-level cameras have much lower resolution than usual optical cameras, which will contaminate the whole range estimation pipeline. Besides, in binocular vision systems, target distance is negatively correlated to disparity, that is, when target moves further, disparity in corresponding image points becomes small, which induces larger error. Due to the differences of starlight, noise and calibrate images, the accidental error of calibration and feature matching also bring the ranging error.

If we can model the variation tendency of ranging error correctly, accurate distances can be obtained through the error correction model. Based on experimental results of multiple groups of data, we find that the ranging error stably increases along with the growth of target distance. By fitting the error, we can compensate the ranging result directly, which avoids all the cumulative error in the entire pipeline. Details are presented as follows.

Suppose the target distance obtained by triangulation and trilateration is z, and its actual range is y. Using collected training images labeled with target actual range (ground truth) and ranging operations mentioned above, we can get a set of scattered points reflecting the relationship of ranging error. With fixed camera position, the error correction model is constructed by fitting $y = f(z)$, with which the ranging error will be dramatically reduced.

$f(z)$ is an inverse of the relation between actual distance and distance estimation. Figure 6 showed the declining trend of distance estimation increase, so the correction value should have an increasing trend. The simplest model is exponential function and power function.

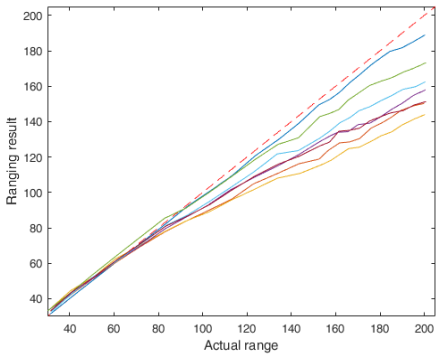


Figure 6. Ranging results of 7 group of experiments. The dotted line is the ideal situation, for the comparation of ranging error.

Based on experimental results, empirical formula can be raised:

$$f(z) = a \cdot e^{b \cdot z} + c \cdot e^{d \cdot z} \tag{10}$$

Equation (10) is a biexponential function, a, b, c, and d are parameters. We use non-linear least squares method to do the fitting. In constrained non-linear optimization algorithm, quadratic model[11] at point z_k is given as:

$$m_k(p) = f(z_k) + \nabla f(z_k)^T p + \frac{1}{2} p^T B_k p \tag{11}$$

B_k is Hessian matrix, which is updated after each iteration, we get search direction p_k by:

$$p_k = -B_k^{-1} \nabla f(z_k) \tag{12}$$

From p_k we get the next point:

$$z_{k+1} = z_k + \alpha_k p_k \tag{13}$$

Besides, three-order polynomial function is also capable of modelling the error:

$$f(z) = a_1 + a_2 \cdot z + a_3 \cdot z^2 + a_4 \cdot z^3 \tag{14}$$

Least square fitting is faster than trusted region reflective algorithm, but the fitting error of three-order polynomial model is a little greater than biexponential model. Equation (10) is utilized in experiments.

3. EXPERIMENTS

3.1 Experimental System and Data

To validate the method, low-light-level CCD camera Watec-910HX with Computar M7528-MP lens is taken as the imaging equipment. The resolution is 720×576 pixels, nominal value of focal length is $75mm$. A sliding rail is used in the experiments to stabilize camera positions. Due to the limitation of our experimental conditions, our method can now be used in remote object ranging with a distance more than 30 meters.

A pair of images used in experiments is shown in Figure 7. The dataset consists of 50 pairs of calibration images, 20 pairs of training images for error correction, 76 pairs of testing images.

Figure 7. Low-light-level images captured for experiments.

3.2 Results

3.2.1 Result of error correction

We choose 25 image pairs from the dataset to do calibration. Pixel focal length has been calibrated as $f_l = (10185.11, 10040.95)$, $f_r = (10259.78, 10182.30)$. To reduce the inaccuracy of calibration, principal points are fixed on $(360, 288)$. The rotation angle from the right camera to the left camera is $r = (0.1971, -3.7358, 1.2843)^T$, and translation vector is $t = (-4001.93, 17.28, 1028.29)^T$.

After ranging the targets on training images using calibration results from above, and paired with ground truth, we can get a set of z-y scattered points. Use biexponential function as error correction model, we fit the points with parameters in 95% confidence:

$$f(z) = 37.2 \cdot e^{0.01229 \cdot z} - 38.83 \cdot e^{-0.01031 \cdot z} \quad (15)$$

$f(z)$ is plotted in Figure 8. Among 30-200 meters, sum squared fitting error is limited to 1.66, which fills the requirements of our task.

3.2.2 *Result of different distances*

With error correction model, we use the entire pipeline to test target range in all 76 image pairs of the dataset, as shown in Figure 9. Compared to the ranging results without error correction (see Figure 6), the ranging error has been dramatically reduced, histogram of the ranging error is shown in Figure 10. As we can see, in most of the test cases, ranging error is less than 1 meter. Mean relative accuracy throughout all tests is 0.18%, which shows that our method is fit for real-life usage.

Compared to conventional approaches conducted on parallel binocular stereo vision system, our method proposes a new pipeline which does not require parallelism of cameras, so it is meaningless to make direct comparisons between the effect of different error corrections. But our error correction method is easier to operate and requires lower computational capacity.

3.2.3 *Result of repeatability test*

We conducted ranging test repeatedly on different distances for 10 times, results are shown in Table 1. The mean relative accuracy decreases as the denominator increases. The absolute amount of ranging error still increases when actual distance becomes larger, but still stays within tolerable limits, since the largest absolute ranging error in this test is $0.656m$, which is obtained in the furthest 200-meter test. The experiment proves that our method produces stable results.

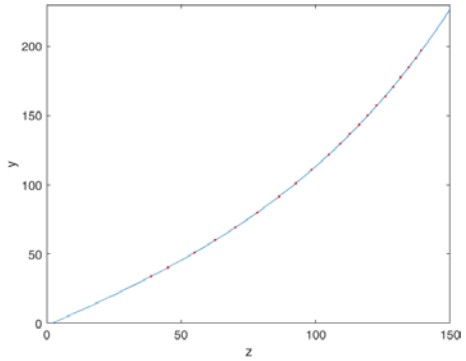


Figure 8. Fitting the ranging error with biexponential model. The dots are ranging results calculated directly by triangulation and trilateration, the y-axis represents ground truth.

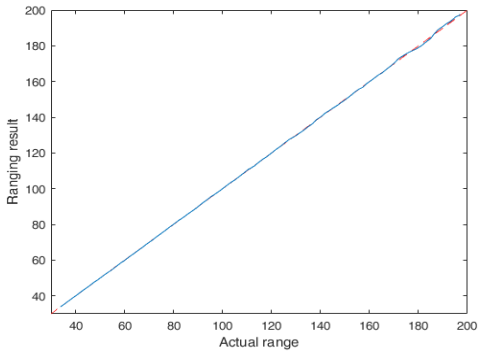


Figure 9. Ranging results with error correction. Line chart represents target distance in test set (76 image pairs). The dotted line is the ideal situation, as a contrast.

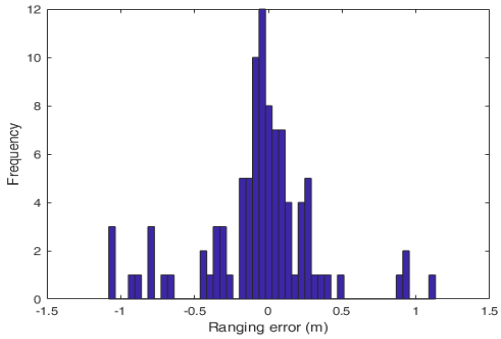


Figure 10. Histogram of ranging error.

3.3 Discussion on System Building

The optical axes of the two cameras are not constrained in parallel locations. That's because we need a baseline long enough for 200-meter ranging, so it would be impractical to maintain the parallel constraint. However, long baseline leads to deficiency in conditions when a target is too close. To avoid gross error, the length of baseline should increase along with distance to the target. Furthermore, distant object becomes blur in images and the feature extraction becomes harder, hence a proper focal length should be chosen. Use a CCD with higher resolution, the ranging accuracy will be improved.

4. CONCLUSIONS

In this paper, a practical low-light-level ranging method based on triangulation and error correction is presented. By using GMS, we have obtained matching points of the target, and the spatial coordinates of the target is calculated by triangulation. Accurate target distance is measured by trilateration along with error correction. Using low cost hardware of low-light-level cameras, the experimental results show that relative accuracy of remote target distance measurement is less than 1%. This method will have a broad application prospects on security surveillance and related operations.

Table 1. Repeatability ranging test on 4 different distances

Actual distance (m)	Mean relative accuracy (%)	Average deviation of ranging error
33.860	0.86%	0.0107
91.570	0.20%	0.0411
150.279	0.12%	0.2206
196.917	0.17%	0.3901

5. REFERENCES

[1] Hartley, R. I. and Zisserman, A. 2000. *Multiple view geometry in computer vision*. Kybernetes, 1333-1341.

[2] Zhang, G. J., Wang, H., Zhao, H. J., Tong, J. L. 1999. Research on structural light 3d vision system. *J. Acta Aeronautica Et Astronautica Sinica*. 20, 4 (Jul. 1999), 365-367.

[3] Huang, G. P., Li, G. Y. and Wang, B. F. 2004. Evolution for monocular vision measurement. *J. Acta Metrologica Sinica*, 25, 4 (Apr. 2004), 314-317.

[4] Jiang, Y. T. 2013. *Binocular Distance Measurement System and Calibration Method*. Master Thesis. Changchun University of Science and Technology.

[5] Tomasi, C. and Manduchi, R. 1998. Bilateral filtering for gray and color images. In *International Conference on Computer Vision* (Mumbai, India, January 4 - 7, 1998), ICCV '98. IEEE, 839-846. DOI=10.1109/ICCV.1998.710815.

[6] Karel, J. Z. 1994. Contrast limited adaptive histogram equalization. *J. Graphics Gems*, 8, 5, 474-485. DOI= https://doi.org/10.1016/B978-0-12-336156-1.50061-6.

[7] Zhang, Z. Y. 2000. A flexible new technique for camera calibration. *J. IEEE Transactions on Pattern Analysis and Machine Intelligence*. 22, 11, 1330-1334. DOI= 10.1109/34.888718.

[8] Rublee, E., Rabaud, V., Konolige, K. and Bradski, G. 2011. ORB: An efficient alternative to SIFT or SURF, In *International Conference on Computer Vision* (Barcelona, Spain, November 06 - 13, 2011), ICCV '11. IEEE, 2564-2571. DOI=10.1109/ICCV.2011.6126544.

[9] Bian, J. W., Lin, W. Y., Matsushita, Y., Yeung, S. K., Nguyen, T. D. and Cheng, M. M. 2017. GMS: Grid-based motion statistics for fast, ultra-robust feature correspondence. In *IEEE Conference on Computer Vision and Pattern Recognition* (Hawaii, USA, July 21 - 26, 2017). CVPR '17.

[10] Hartley, R. I. and Sturm, P. Triangulation. 1997. *J. Computer Vision and Image Understanding*, 68, 2, 146-157.

[11] Vanderbei, R. J., Meketon, M. S. and Freedman, B. A. 1986. A modification of karmarkar's linear programming algorithm. *J. Algorithmica*, 1, 1-4, 395-407.