Research Article

An Analytical Measuring Rectification Algorithm of Monocular Systems in Dynamic Environment

Deshi Li[1,2] and Xiaoliang Wang[1]

[1]*Electronic Information School, Wuhan University, Wuhan, Hubei 430072, China*
[2]*Collaborative Innovation Center of Geospatial Technology, 129 Luoyu Road, Wuhan 430079, China*

Correspondence should be addressed to Xiaoliang Wang; xiaoliangwang@whu.edu.cn

Range estimation is crucial for maintaining a safe distance, in particular for vision navigation and localization. Monocular autonomous vehicles are appropriate for outdoor environment due to their mobility and operability. However, accurate range estimation using vision system is challenging because of the nonholonomic dynamics and susceptibility of vehicles. In this paper, a measuring rectification algorithm for range estimation under shaking conditions is designed. The proposed method focuses on how to estimate range using monocular vision when a shake occurs and the algorithm only requires the pose variations of the camera to be acquired. Simultaneously, it solves the problem of how to assimilate results from different kinds of sensors. To eliminate measuring errors by shakes, we establish a pose-range variation model. Afterwards, the algebraic relation between distance increment and a camera's poses variation is formulated. The pose variations are presented in the form of roll, pitch, and yaw angle changes to evaluate the pixel coordinate incensement. To demonstrate the superiority of our proposed algorithm, the approach is validated in a laboratory environment using Pioneer 3-DX robots. The experimental results demonstrate that the proposed approach improves in the range accuracy significantly.

1. Introduction

The applications of mobile robots for observation and rescue missions have received an increasing attention in recent years. Current advances in sensing and computing promote mobile robots as a suitable option in occasions such as search and rescue, SLAM (Simultaneous Localization and Mapping), automatic navigation, and target detection. For mobile robots, retrieving their position is one of the important issues. In recent years, to solve this problem, vision sensors have attracted a lot of attention because vision sensors are relatively inexpensive and compact with low power consumption. Furthermore, methods using version sensors can localize the robot in various environments where it is difficult for general localization methods like wheel odometry and GPS. If localization can be performed only using image information, a robot's flexibility will be improved remarkably.

However, the precondition to a successful intelligent robot system is the exact perception of surroundings, where the range and azimuth information of targets around play an important role. The range estimation algorithms using vision sensors are known as VO (visual odometry).

Approaches for range estimation can mainly be divided into three categories: radar-based, laser-based, and vision-based. As a typical paradigm of noncontact approaches, ultrasonic sensors have the advantages of time efficiency and measurement accuracy. However, it is arduous to detect those objects with small surfaces or situated at a wide angle related to the ultrasonic sensor(s). Among all perception sensors, computer visions have an added advantage of acquiring large amount of information at a lower cost. The vision-based method can solve both range and azimuth estimation problems using only the acquired image themselves. There has been much interest in research on object detection by stereo camera [1–4], but the monocular camera is still strongly advantageous for its large sensing area, low cost, and easy installation.

To achieve precise VO in outdoor environments, some problems remain to be solved. In this work, the problems of VO on bumpy courses which exist mostly in outdoor

environments are considered, where the bumpy courses mean that the environments include rough roads on which the VO accuracy is dynamically affected by the pose change of vision sensors. Furthermore, if precise VO is realized in the environments including rough roads, we believe that it can be utilized in any outdoor environments.

In the researches of intelligent unmanned vehicle systems, computer vision generally adopts the methods of imaging processing algorithms. In those works, the image features are extracted, along with the model of the ambient environment, for vehicle localization and obstacle avoidance. Range and azimuth information are then refined from the above model using vision system. It is unrealistic to assume that the road is absolutely flat in the process. This paper concentrates on the dynamic measurement rectification problem in which the camera pose changes abruptly. This approach is particularly suitable for applications such as navigating the autonomous vehicles running on rough terrains. Pose variations are firstly measured by a three-axis angle sensor and sequentially applied to calculate the distance offsets using the proposed range-pixel model. Although monocular visual odometry has an advantage in wide FOV (field of view), factors such as lighting conditions, shadows, and random noise would unavoidably decrease the measurement precision, which are induced from both human limitations and sensor characteristics. On the contrary, noncontact sensors such as sonar are typically not susceptible to those external conditions which would infect the result accuracy. Nevertheless, one main defect is the existence of inherent blind areas. In the view of these possible advantages and corresponding limitations, sensor assimilation technique based on OI (Optimal Interpolation) method is employed. The main contributions of this paper are summarized as follows.

(1) The relation between range increment and camera's pose variation has been formulated, based on which a feasible data rectification algorithm has been designed to modify the metrical results. To the best of our knowledge, it is the first work to solve range estimation problem under camera shaking conditions.

(2) An improved estimation mechanism of range information in OI model has been developed, which enhances adaptability and accuracy of multisensor measuring system.

(3) Experiments on mobile robots and analytical results have been demonstrated.

The rest of this paper is organized as follows. The following section will provide some background and a more detailed literature review. Section 3 defines the problems and related literatures. Section 4 details the proposed approach for measurement rectification and sensor fusion. Finally, experiment results and conclusions are given in Sections 5 and 6.

2. Related Works

Visual distance estimation is a specialized set of approaches which focus on real-time and accurate image capture followed by range information acquirement. Several of these mechanisms have been developed as foundational elements of 3D reconstruction, simultaneous localization, and map building.

Some basic algorithms as well as their improvements for range estimation have been developed: epipolar constraint model [2], defocusing method [3–5], coordinate mapping scheme [6, 7], and camera movement approach [1, 8]. Katsuyuki et al. proposed a coupled estimation of unknown vehicle width and following distance by sequential Bayesian estimation. The method can run in real-time and produce highly accurate estimation of the following distance under a precondition that no camera shaking happens.

Those proposed methods can be divided into two categories: monocular and stereo system. Monocular approaches involve a single no sophisticated camera that compute the pixel size or coordinates which are used for range estimation. Examples of these are studied in [9]. Stereo vision approaches can provide much higher accuracy than monocular, but they have small field of view and high operational complexity. Several intelligent and operable algorithms [10, 11] fall into this category.

Monocular and stereo vision approaches have advantages in different aspects. Monocular approaches are usually easy to be implemented and have optimal view scope. Meanwhile, they require much lower cost compared to the former. Stereo vision methods, in contrast, have a good performance in accuracy due to the subpixel synthetical localization technology, while their biggest drawback lies in the complicated operations and high computational complexity, especially during the calibration process.

Among these emerged researches, most work assumes the camera pose is fixed [3, 6, 9, 10, 12–14]. Some notable exceptions, which have similarity to the present work, are as follows. (1) Guo et al. [15] put forward a parallel constraint method based on the two lane boundaries. (2) Vehicles are equipped with an angle sensor to accurately acquire the pitching angle of the camera in [13, 16, 17], where the author proposed an improved algorithm in angle calculation by using a function of the angles representing the two parallel lane lines.

Some other approaches have also been proposed. Typical paradigms are as follows. Han et al. [18] devise a feature point based method for monocular measurement, but they hinder the real-time implementation. Malis and Rives [19] design a hybrid algorithm to minimize the token relative displacements between two frames and then estimate the image-space distance.

3. Problem Formulation

In Figure 1 suppose $\mathbf{P}'$ is a point in the image plane of the camera in a pose of $\mathbf{P_s}$, and suppose that we have an estimate of the pose of the camera by a three-axis gyroscope. From this information a standard ground-constrained model [18] can be used to estimate the position of $\mathbf{P}$ in the world coordinate. If the camera's pose suddenly changes to $\mathbf{P_f}$, we can use this information to project point $\mathbf{P}$ into the camera's image plane,
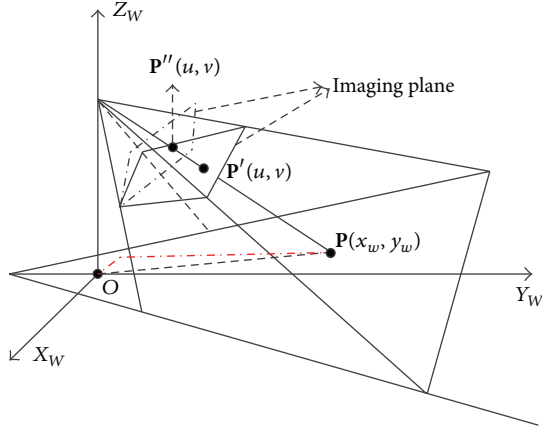
FIGURE 1: The imaging geometry of the observation model.

FIGURE 2: The world and camera coordinate systems.

which obtains a second point $\mathbf{P}''$. Now assuming that the pose measurement is reasonably accurate and that the position estimate algorithm works well, the problem is to estimate $\mathbf{P}$ utilizing measurements including $\mathbf{P}'$, $\mathbf{P}''$, and pose variation of the camera.

The initial and final poses of the camera are denoted by $\mathbf{P_s} = (r_s, p_s, y_s)$ and $\mathbf{P_f} = (r_f, p_f, y_f)$, respectively, where r_i, p_i, and y_i ($i = s, f$) stand for the initial roll, pitch, and yaw angles. Although the actual relative distance from the optical center to the target changes slightly, the measured results deviate from the truth significantly. This is mainly because of the nonlinear mapping between pixel coordinates and corresponding distance values. The problem is to correct the actual measurements to be close to the truth by eliminating the pose perturbation of the camera.

4. Data Rectification Algorithm

In this section we describe our approach to the problem of monocular vision-based measurement rectification. Since a robot's trajectory is most conveniently described in a world coordinate system, while the target on the ground is generally described by its camera coordinate system, we start with a preview of these two coordinate systems. To model the problem in a general geometrodynamical architecture, the algebraic relation between the camera pose displacement and the displacement of measuring distance is derived.

4.1. World and Camera Coordinates System. Assume that XYZ and $X'Y'Z'$ are, respectively, a world coordinate system and the camera's coordinate system as shown in Figure 2. The coordinates of a point $\mathbf{P}$ under these two coordinate systems are transformed by

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{R} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + \mathbf{T}, \tag{1}$$

where (X, Y, Z) and (X', Y', Z') are point $\mathbf{P}$ coordinates in the world and robot camera coordinate system. Moreover,
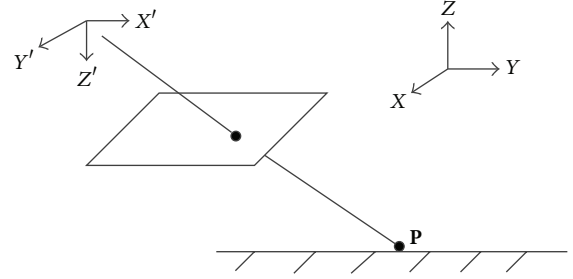
$\mathbf{R}$ and $\mathbf{T}$ are, respectively, the rotation and translation from the camera to the world's coordinate system, which determine the position and orientation of the camera in the world coordinate system. Furthermore, for a 3D point $\mathbf{P}$ in the FOV of the camera, its image coordinates are given by the projection equation as follows:

$$[u \ \ v]^T = \left[\frac{X'}{Z'} \ \ \frac{Y'}{Z'} \right]^T \cdot f, \tag{2}$$

where (u, v) are the coordinates of $\mathbf{P}$ in the image coordinate system and f is the camera's focal length.

4.2. Chebyshev Best Uniform Approximation Rectification Algorithm. The distance-orientation information between targets and the camera can be derived from corresponding pixel coordinate in the image [20, 21]. It is found that the ratio of image pixel motion to the camera rotation angles varies nonlinearly along the main optical axis. The main idea of the designed algorithm is to piecewise linearize the nonlinear rate and then calculate the rate of change with respect to rotation angles as well as the measured distance. Equation (3) presents the rotation matrix in 3D space:

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{bmatrix}$$
$$\cdot \begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3}$$

Variations of pixel coordinates are associated with the world coordinates by a rotation matrix whose parameters are attitude angles of the camera, which is described by

$$\Delta u_{k+1} = F_x \cdot \left(\frac{R_{11}^{k+1} x_w + R_{12}^{k+1} y_w}{R_{31}^{k+1} x_w + R_{32}^{k+1} y_w} - \frac{R_{11}^k x_w + R_{12}^k y_w}{R_{31}^k x_w + R_{32}^k y_w} \right),$$
$$\Delta v_{k+1} = F_y \cdot \left(\frac{R_{21}^{k+1} x_w + R_{22}^{k+1} y_w}{R_{31}^{k+1} x_w + R_{32}^{k+1} y_w} - \frac{R_{21}^k x_w + R_{22}^k y_w}{R_{31}^k x_w + R_{32}^k y_w} \right), \tag{4}$$

where inner parameters F_x and F_y are only determined by the CCD structure itself.

For the convenience of discussion, we assume that the camera poses change mainly along the yaw angle direction.

Denote $\Psi(\theta, \psi, \varphi) = (R_{11}x_w + R_{12}y_w)/(R_{31}x_w + R_{32}y_w)$; using (3) and (4), we obtain

$$\Psi(\theta, \psi, \varphi) = \cos \partial + \sin \partial \cot \beta, \tag{5}$$

where

$$\partial = \arctan\left(\frac{x_w}{y_w}\right) + \arctan\left(\frac{y_w}{x_w}\right),$$

$$\beta = \varphi - \arctan\left(\frac{y_w}{x_w}\right). \tag{6}$$

Substituting (3) in (5) results in

$$\Psi(\theta, \psi, \varphi) = \cot\left(\varphi - \arctan\left(\frac{y_w}{x_w}\right)\right). \tag{7}$$

From Figure 3, we can see that the slope of curves tends to be constant within a sliding interval of independent variable. This interval becomes smaller when the ratio of x_w to y_w increases. The Chebyshev approximation method has the characteristics of uniform approximation on selected closed-interval. Inspired by this, the nonlinear rate can be approximated by linear polynomial and the deviation caused by poses change of a camera can be effectively compensated. The second derivative is taken as

$$\Psi''(\varphi) = 2\cot\left(\varphi - \arctan\left(\frac{y_w}{x_w}\right)\right)$$

$$\cdot \left(\cot\left(\varphi - \arctan\left(\frac{y_w}{x_w}\right)\right)^2 + 1\right). \tag{8}$$

Considering the yaw angle variations of a PTZ (Pan/Tilt/Zoom) camera caused by uneven pavement during practical robot motion, a closed subinterval $[a, b]$ ($[a, b] \subseteq [0 \sim \pi/12]$) is chosen for further deduction. Since (8) is a continuous function and keeps consistency in sign, the best consistent approximation method can be used.

Denote $a_1^* = (\Psi(b) - \Psi(a))/(b - a)$, using this to acquire normal equation of approximation:

$$\Psi'(\varphi) = a_1^*. \tag{9}$$

Set the solution of (9) as x_Δ. Then the approximation equation is written as follows:

$$p^*(x) = \frac{f(a) + f(x_\Delta)}{2} + a_1^* \cdot \left(x - \frac{a + x_\Delta}{2}\right). \tag{10}$$

We explore the slope of line after linear approximation to study the function of different ratios of x_w to y_w. Results show that the slope converges to its limit uniformly. Moreover, this constant value is irrelevant to the ratio above:
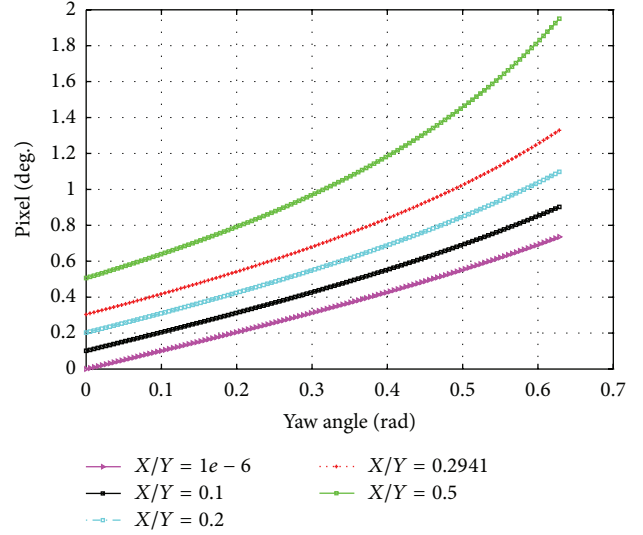
$$\kappa = \frac{\Psi(b) - \Psi(a)}{b - a}. \tag{11}$$

FIGURE 3: Pixel coordinates along U-axis vary with the yaw angles: a set of y_w/x_w values are considered. (x_w, y_w) is the coordinate of **P** in the world coordinate system.

Substituting (7) in (11) results in

$$\kappa = \frac{1}{b - a} \cdot \frac{1 + (y/x)^2}{(\tan(b) - y/x)(\tan(b) - y/x)}$$

$$\cdot (\tan(a) - \tan(b)). \tag{12}$$

To demonstrate the convergence of κ, we have also analyzed the limit value given by

$$\lim_{y/x \to \infty} k = \frac{\tan(a) - \tan(b)}{b - a}. \tag{13}$$

Solid curves in Figure 4 are the results of actual slope and linear approximation, respectively. These two curves coincide with each other well after a translation operation. This indicates a high accuracy in slope using linear approximation. Figure 5 manifests the convergence of slope related to a metric of y_w/x_w, which is in good agreement with experimental results. Another important property that should be noted is that the function value rapidly reaches convergence after a dramatic increase; that is, the measured range would vary with the metric y_w/x_w with high nonlinearity. This also implies that the measurement should be conducted on the smooth interval to reduce the errors caused by camera shakings. On the other hand, it is impossible to compensate the deviations when the metric is too small.

4.3. Sonar and Camera Data Assimilation Model.

The Optimal Interpolation Algorithm is derived to generate the least squares results for vectors of observations and background fields assuming "a priori" known statistical models for the background error covariance. The Optimal Interpolation Technique, based on the minimization of variance estimation, plays an important role in data assimilation. It uses several different real-world observations to produce a corrected output, which is closer to the truth.
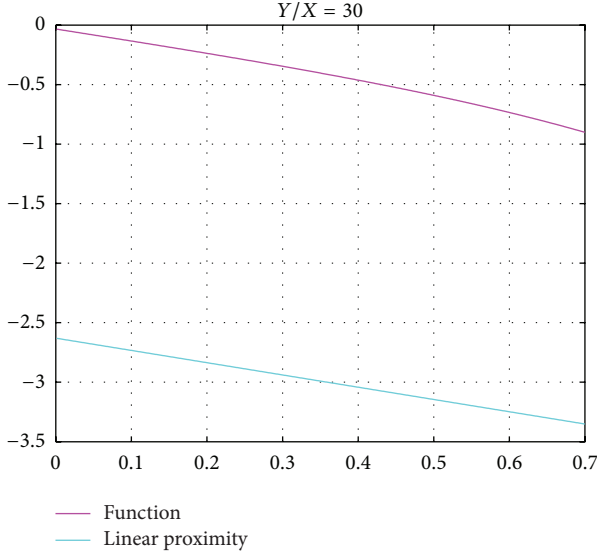
FIGURE 4: Pixel values vary with the yaw angle: our approach can successfully approximate the actual changes.
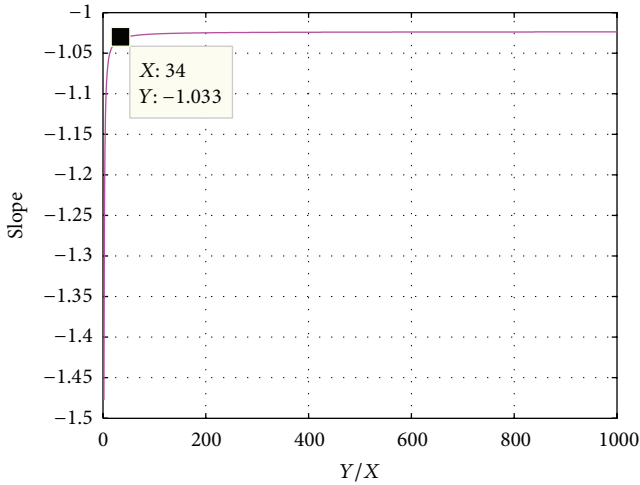

FIGURE 5: Slope converges to a stable value.

The motivation of the proposed method comes from the similar characteristic and phenomena between a camera and sonar measurement system and an OI algorithm. First, a camera and sonar system can be considered to be an OI system that produces an optimal output through several groups of observations. Second, the OI algorithm has the dimension-extensible and loose a priori characteristics that are attractive for camera and sonar measurement system.

The following are given:

(i) A background field x_b available in two or three dimensions.

(ii) A set of p observations x_{obs} available at irregular positions.

The optimal estimation is described by

$$x_a = x_b + W\left(x_{obs} - x_b\right). \tag{14}$$

TABLE 1: Initial state of camera before calibration.

Rotation angle	Pitch angle	Zoom
0	$-15°$	0

TABLE 2: Calibration results.

f_x	f_y	U_0	V_0
731.88	738.65	339.28	249.7

The errors are given by

$$\varepsilon^a = x_a - x_t,$$

$$\varepsilon^b = x_b - x_t, \tag{15}$$

$$\varepsilon^{obs} = x_{obs} - x_t.$$

The optimal weight is then as follows:

$$W = \frac{\left(\sigma^b\right)^2}{\left(\sigma^{obs}\right)^2 + \left(\sigma^b\right)^2}, \tag{16}$$

where σ^a and σ^b represent the mean value of ε^a and ε^b. As data from camera and sonar are unrelated, it is assumed that $E(\varepsilon^b \varepsilon^{obs}) = 0$.

5. Evaluations and Analysis

In this section, we present the results of a set of physical experiments to demonstrate the performance of the proposed algorithm in Section 4. To validate the effectiveness of the proposed data rectification algorithm, we compared the results before and after a pose change with the truth. Moreover, we have conducted a set of experiments under different initial poses of a camera to testify the robustness of this method. Besides, comparative experiments have been designed to show the validity of the data assimilation approach.

Autonomous vehicles can be modeled as mobile robots and then we use the mobile robot Pioneer 3-DX (Figure 6) mounted with a camera for experiments. To prepare for experiments, the PTZ camera is firstly calibrated.

5.1. Camera Calibration. Grid size of the calibration board in experiments is 30 mm ∗ 30 mm. Picture resolution of VCC50I is fixed as 640 ∗ 480. To ensure error balance and calibration accuracy, a group of calibrated images containing four images from various poses are collected at a distance interval of 10 cm. Calibration distance ranges from 1500 mm to 4000 mm. Considering effects of pitch and rotation angles as well as the zoom value, the camera state during calibration is fixed as given in Table 1. Internal parameters are listed in Table 2, which are crucial to distance measurement.

5.2. Performance Evaluation of the Data Rectification Algorithm. Angle variations are acquired by a three-axis angle sensor, which act as the input of the rectification algorithm.

Figure 6: Pioneer 3-DX robot equipped with a Canon-VCC50I PTZ Camera.

Table 3: Different conditions set for measurement.

ψ	θ	φ
2.13	7.18	0.05
4.89	4.84	2.81
5.68	4.05	3.59
2.93	6.42	2.77
4.07	5.01	1.10

The module (MPU6050) has advantages in low temperature dependency, high resolution, and low noise. Due to these advantages, this module is chosen as a tool to measure the Euler angles.

To validate the robustness of the algorithm, a target is set at different positions randomly. For each metric y/x, results under a set of camera poses are analyzed (yaw, pitch, and roll angles are random for each metric as set in Table 3). Initial readings of angle sensor are as follows: θ (pitch): 9.65, φ (yaw): $-0.27°$, and ψ (roll): $-0.94°$. The target position tuples are set as {(1500, 100), (2000, 100), (2500, 100), (3000, 100)}.

Figure 7(a) shows measurements from the initial camera state. Compared with the truth, it shows bias along both horizontal and vertical directions.

The deviations of VO results caused by the camera motions are rectified independently. Based on the analysis of the range model in Section 3, the pitch angle is an independent variable of distance function. Therefore, we recalculate the pitch angle instead of inversion operations on pixel coordinates. Figures 7(b)–7(f) show results before and after rectification. In Figure 7(e), the distance error along the optical axis is almost as high as 50% using direct measurement. However, this value decreases to be only 6% using the proposed algorithm. We can also see that the least improvement in accuracy in this direction is 10% as shown in

Figure 7(c). Accordingly, much more remarkable effect can be seen from the results along the direction perpendicular to the optical axis. In the worst case as shown in Figure 7(f), the measured distance along X-axis is rectified from the measured -780 mm to the final 98 mm. The percentage gains of measuring precision approach 878%. Even in a general situation, this percentage can be close to 35% as demonstrated in Figure 7(c). These figures also show that the range deviation becomes larger as the distance along the optical axis direction increases. This is mainly because the ratio of physical distance to pixel unit increases along the optical axis.

5.3. Data Assimilation Evaluation. For generality, the assimilation results under different metric size (i.e., manipulating y/x) with a fixed camera pose are demonstrated. Data from sonar sensors are set as the background field value and those from camera are set as the observation field value. In Figure 8(a), measurement results at some positions are missing, which indicate that blind zones exist when sonar system only is adopted. Range data in Figure 8(b) are the results of sensor assimilation. It demonstrates the accuracy improvement from both X-axis and Y-axis compared with measurements solely from the vision system and sonar sensors. Assimilated results proved to be as much as 25 percent accurate along X-axis and 9 percent accurate along Y-axis compared to those acquired using a single type of sensor. This is mainly because new information is brought in to compensate the output from a single measuring system, that is, the wild FOV of camera and the high measurement accuracy of sonar sensors.

6. Conclusions

In this paper, we have proposed an analytical measuring rectification algorithm for monocular range estimation under camera shaking conditions. Specifically, we have established a pose-range model and then the algebraic relation between

$\theta = 9.65°$ $\phi = -0.27°$ $\psi = -0.94°$

(a)

$\theta = 7.18°$ $\phi = 0.05°$ $\psi = 2.13°$

- * - Camera measured value
- o - True value
- △ - Rectified value

(b)

$\theta = 4.84°$ $\phi = 2.81°$ $\psi = 4.89°$

- * - Camera measured value
- o - True value
- △ - Rectified value

(c)

$\theta = 4.05°$ $\phi = 3.59°$ $\psi = 5.68°$

- * - Camera measured value
- o - True value
- △ - Rectified value

(d)

$\theta = 6.42°$ $\phi = 2.77°$ $\psi = 2.93°$

- * - Camera measured value
- o - True value
- △ - Rectified value

(e)

$\theta = 5.01°$ $\phi = 1.10°$ $\psi = 4.07°$

- * - Camera measured value
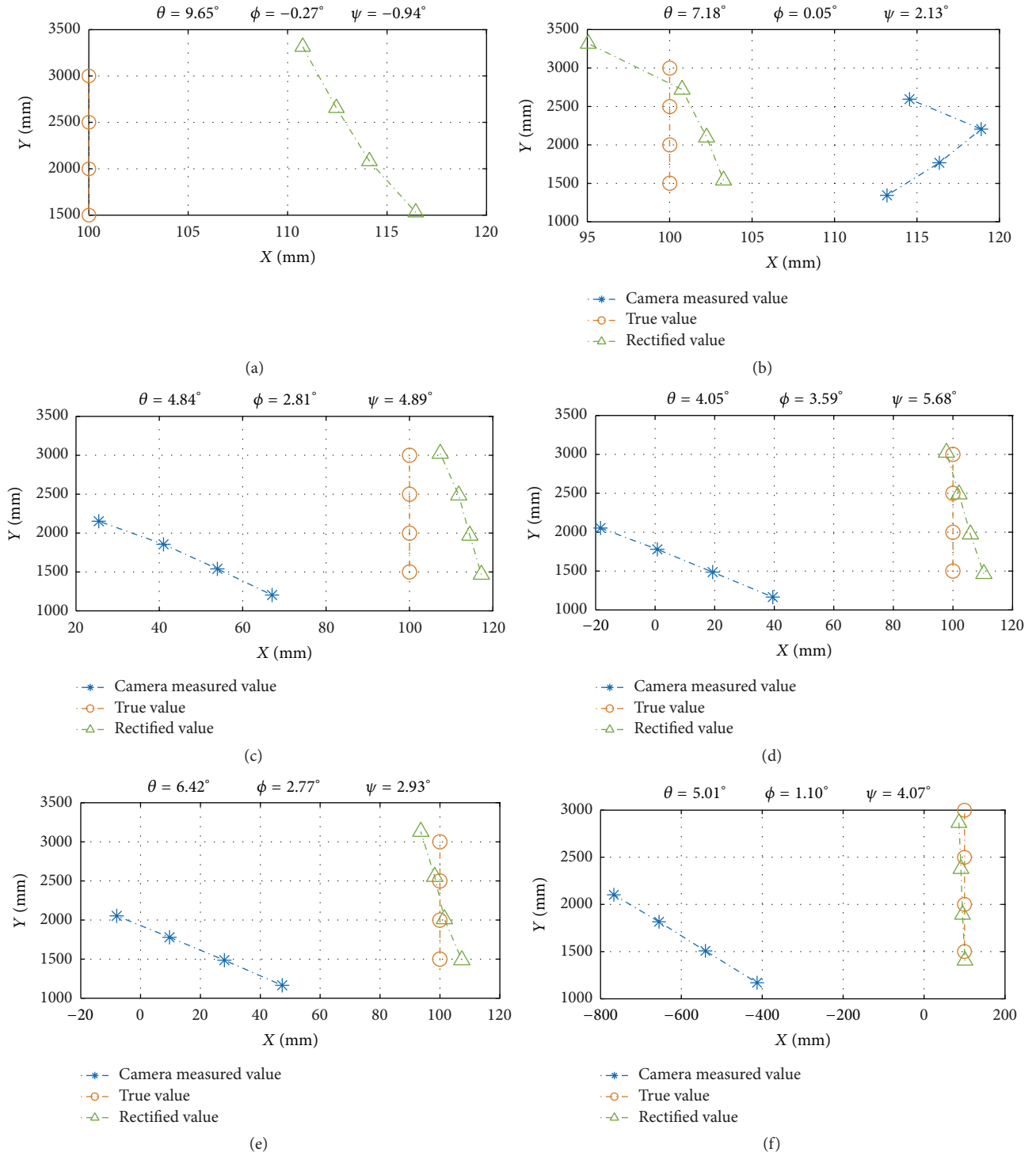- o - True value
- △ - Rectified value

(f)

FIGURE 7: Comparisons between results after and before rectification: a set of camera poses are set to validate the robustness of designed algorithm.

distance increment and a camera's poses variation has been formulated. We have also designed a data assimilation system to provide reliable range information using different types of transducer systems. Physical experiments are conducted to validate the effectiveness and robustness of the proposed algorithm. For future work we will try to implement our algorithms on the multiple robots formations as well as swarm coordination applications.

(a) Results before assimilation
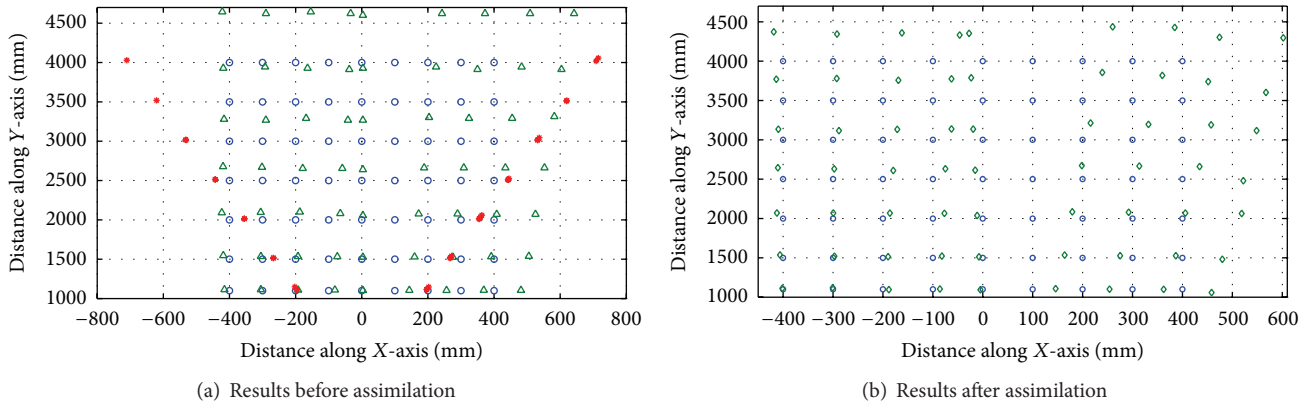
(b) Results after assimilation

FIGURE 8: Comparisons between results before and after assimilation. (a) Original data from sensors of sonar and camera. (b) Results of assimilation using the Optimal Interpolation Algorithm. (Results of camera are represented as Δ and those of sonar are described as $*$. The symbol $\circ$ stands for truth values and $\diamond$ are assimilated results.)

Competing Interests

The authors declare that they have no competing interests.

References

[1] M. Mikhalsky and J. Sitte, "Active motion vision for distance estimation in autonomous mobile robots," in *Proceedings of the IEEE Conference on Robotics, Automation and Mechatronics*, vol. 2, pp. 1060–1065, December 2004.

[2] N. Yamaguti, S. Oe, and K. Terada, "A method of distance measurement by using monocular camera," in *Proceedings of the 36th SICE Annual Conference (SICE '97)*, International Session Papers, pp. 1255–1260, IEEE, Tokushima, Japan, July 1997.

[3] X. Liu, T. Qin, W. Chen, and X. Yin, "Real-time distance measurement using a modified camera," in *Proceedings of the IEEE Sensors Applications Symposium (SAS '10)*, pp. 54–58, IEEE, Limerick, Ireland, February 2010.

[4] U. Mudenagudi and S. Chaudhuri, "Depth estimation using defocused stereo image pairs," in *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV '99)*, vol. 1, pp. 483–488, IEEE, September 1999.

[5] A. N. Rajagopalan, S. Chaudhuri, and U. Mudenagudi, "Depth estimation and image restoration using defocused stereo pairs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1521–1525, 2004.

[6] M. N. A. Wahab, N. Sivadev, and K. Sundaraj, "Target distance estimation using monocular vision system for mobile robot," in *Proceedings of the 2nd IEEE International Conference on Open Systems (ICOS '11)*, pp. 11–15, IEEE, Langkawi, Malaysia, September 2011.

[7] K. Umeda and T. Takahashi, "Subpixel stereo method: a new methodology of stereo vision," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '00)*, vol. 4, pp. 3215–3220, San Francisco, Calif, USA, April 2000.

[8] T.-H. Wang, C.-C. Hsu, C.-C. Chen, C.-W. Huang, and Y.-C. Lu, "Three-dimensional measurement of a remote object with a single CCD camera," in *Proceedings of the 4th International Conference on Autonomous Robots and Agents (ICARA '09)*, pp. 187–192, IEEE, Wellington, New Zealand, February 2009.

[9] K. Nakamura, K. Ishigaki, T. Ogata, and S. Muramatsu, "Real-time monocular ranging by Bayesian triangulation," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV '13)*, pp. 1368–1373, IEEE, Gold Coast, Australia, June 2013.

[10] A. Tsalatsanis, K. Valavanis, and N. Tsourveloudis, "Mobile robot navigation using sonar and range measurements from uncalibrated cameras," *Journal of Intelligent and Robotic Systems*, vol. 48, no. 2, pp. 253–284, 2007.

[11] V. Tucakov, M. Sahota, D. Murray et al., "A stereoscopic visually guided mobile robot," in *Proceedings of the Hawaii International Conference on Systems Sciences*, Maui, Hawaii, USA, January 1997.

[12] B. Fidan, V. Gazi, S. Zhai, N. Cen, and E. Karatas, "Single-view distance-estimation-based formation control of robotic swarms," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 12, pp. 5781–5791, 2013.

[13] X. Li and L. Wang, "A monocular distance estimation method used in video sequence," in *Proceedings of the IEEE International Conference on Information and Automation (ICIA '12)*, pp. 390–394, Shenyang, China, June 2012.

[14] J. Yu, J. Song, X. Sun, and T. Zhang, "Design and experiment of distance measuring system with single camera for picking robot," in *Proceedings of the IEEE World Automation Congress (WAC '10)*, pp. 445–448, Kobe, Japan, September 2010.

[15] L. Guo, Y. Xu, K. Li et al., "Study on real-time distance detection based on monocular vision technique," *Journal of Image and Graphics*, vol. 11, no. 1, pp. 74–81, 2006.

[16] S. Derrouich, K. Izumida, and K. Shiiya, "A combination of monocular CCD camera and inertial-sensor for range estimation," in *Proceedings of the 28th Annual Conference of the IEEE Industrial Electronics Society (IECON '02)*, pp. 2191–2196, IEEE, November 2002.

[17] D. Wang, W. Chen, Z. Zhao, and T. C. Ng, "Inter-vehicle separation measurement using monocular vision," in *Proceedings of the 7th IEEE Conference on Industrial Electronics and Applications (ICIEA '12)*, pp. 258–263, Singapore, July 2012.

[18] Y. X. Han, Z. S. Zhang, and M. Dai, "Monocular vision system for distance measurement based on feature points," *Optics and Precision Engineering*, vol. 19, no. 5, pp. 1110–1117, 2011.

[19] E. Malis and P. Rives, "Robustness of image-based visual servoing with respect to depth distribution errors," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '03)*, vol. 1, pp. 1056–1061, IEEE, Taipei, Taiwan, September 2003.

[20] R. F. Stengel, *Optimal Control and Estimation*, Dover, New York, NY, USA, 1994.

[21] J. Kim, Y. Kim, and S. Kim, "An accurate localization for mobile robot using extended Kalman filter and sensor fusion," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '08)*, pp. 2928–2933, IEEE World Congress on Computational Intelligence, Hong Kong, June 2008.