Error modelling of depth estimation based on simplified stereo vision for mobile robots

Bo Jin*, Lijun Zhao, Shiqiang Zhu

The State Key Lab of Fluid Power Transmission and Control, Zhejiang University, Hangzhou, 310027, China

Received 1 June 2014, www.cmnt.lv

Abstract

Depth estimation is the precondition in obstacle avoidance for mobile robots. To improve the obstacle detecting effectiveness and quickness in poor-textured backgrounds, we used the centroid abscissa difference of corresponding obstacle region in image pairs as parallax to estimate obstacle depth. The error of parallax and depth were studied analytically and numerically. Wood blocks of different shapes and sizes were used for demonstrating the relationship between estimated depth and actual depth. A quadratic function model was obtained after experiments. Although the depth estimation error was relatively higher compared to conventional grayscale correlation-based method, the proposed method was expected to satisfy the accuracy requirement of depth estimation for common mobile robots.

Keywords: error model, depth estimation, stereo vision, radial distortion

1 Introduction

Depth estimation is prerequisite to obstacle avoidance for mobile robots and various sensors can be adopted [1, 2]. Being lightweight, power-efficient and inexpensive, stereo cameras are preferred in range finding [3, 4]. However, cameras need accurate calibration and algorithms rely seriously on scene textures. To guarantee estimation accuracy, image rectification is inevitable. Meanwhile, stereo matching is needed which uses similarity functions including region descriptors [5, 6] and feature descriptors [7, 8]. These procedures lead to heavy calculation burden. Moreover, high mismatching rate happen to applications in poor-textured scenes. In this paper, we presented a strategy for depth estimation in poor-textured scenes. Conventional distortion rectification was avoided and density-based matching was substituted with calculating the centroid abscissa difference of corresponding obstacle regions in image pairs. That difference was taken as parallax to estimate obstacle depth.

Zou analysed the influence on depth estimation by hardware system error, camera calibration error, feature extraction error and stereo matching error [9]. Rodriguez analysed the quantization error of stereo vision system and presented the probability density functions [10]. Fooladgra gave a geometrical approach to estimate the amount of localization error [11]. However, the studies mainly focused on theoretical analysis. Some models help to minimize depth estimation error but increase computational burden, so they are more suitable for preliminary theoretical design rather than real-time application. Rodriguez indicated that acceptable error should be decided before selecting relevant parameters [10]. Llorca suggested that a trade-off should be reached between accurate estimation and other parameters [12]. So we suggested estimation strategy be based on acceptable error in specific applications. For mobile robots, less depth estimation accuracy is acceptable in return for real-time performance. The lost accuracy can be balanced by setting safe margin for obstacle areas referencing the strategies of UAV obstacle avoidance and route planning which represents the obstacle areas as circles [13]. As to the errors of parallax and depth estimation in our strategy, analytical and numerical studies were given in Sections 2 and 3.

2 Depth estimation error analysis

The strategy we proposed for obstacle depth estimation included three steps. First, image pairs were block-segmented in space domain. Then feature vector of candidate obstacle areas were extracted. Finally, the centroid abscissa difference of corresponding obstacle areas was taken as parallax for depth estimation. There were several studies on image block segmentation [14]. Meanwhile, the components of feature vector are mainly coordinates easy to be extracted. So we mainly focused on modelling of obstacle depth estimation error, assuming that images are segmented correctly and the extracted centroids coincide with their actual position.

Considering point $P(x_w, y_w, z_w)$ in the world coordinate system (Figure 1), if the projective points of P in left and right image planes are $P_l(x_l, y_l)$ and $P_r(x_r, y_r)$ in each image coordinate system, then the theoretical depth of P can be expressed based on triangulation method:

* ***Corresponding author*** e-mail: bjin@zju.edu.cn

$$z = \frac{B \cdot f}{\lambda \cdot d}, \tag{1}$$

where B is baseline length and f is camera lens focal length. The parallax of corresponding points is denoted as d. The size of pixel is λ. Assuming $x_l > x_r$, then we have:

$$d = x_l - x_r. \tag{2}$$

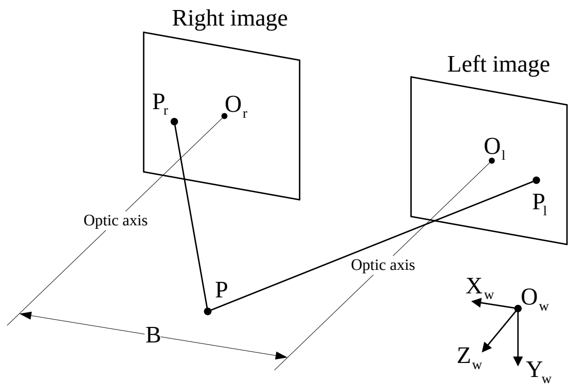


FIGURE 1 The stereo vision system model

The pixel size is given in camera manual. The focal length can be calculated by calibration. The measurement accuracy of B is easy to be guaranteed. So we take only the error of parallax d into account while analysing the error of depth z.

The error of parallax caused by lens distortion is denoted as Δd, then the relationship between measured parallax d' and ideal parallax d is given by:

$$d' = d + \Delta d. \tag{3}$$

Thus, the estimated depth can be calculated as:

$$z' = \frac{B \cdot f}{\lambda \cdot d'}. \tag{4}$$

Using Equations (1) and (3), the error of depth estimation can be written as:

$$\begin{aligned} \Delta z = z - z' &= \frac{B \cdot f}{\lambda \cdot d} - \frac{B \cdot f}{\lambda \cdot d'} = \\ \frac{B \cdot f}{\lambda \cdot d} - \frac{B \cdot f}{\lambda \cdot (d + \Delta d)} &= z^2 \frac{\lambda}{B \cdot f} \cdot \frac{\Delta d}{1 + \Delta d / d}. \end{aligned} \tag{5}$$

Based on Equation (5), we expected that within certain range $(z_{\min}, z_{\max})$ for z, if:

$$g(\Delta d, d) = \frac{\lambda}{B \cdot f} \cdot \frac{\Delta d}{1 + \Delta d / d} \approx C_e(\text{constant}),$$

the following relation will become true:

$$\Delta z = C_e \cdot z^2. \tag{6}$$

Among common errors [9], calibration error and matching error are not involved in this paper. Ignoring the hardware system error and feature extraction error, we analyse the parallax error caused by camera lens distortion.

Camera lens distortions mainly include radial distortion, decentring distortion and prism distortion [15]. We considered only the radial distortion here because it dominates in all the distortions.

In image coordinate systems, the projective points of optical centre for left and right camera lens are denoted as $O_l(x_l, y_l)$ and $O_r(x_r, y_r)$, and we have the following commonly used equations:

$$\begin{aligned} &x_l' - u_l = (x_l - u_l) + \\ &k_l(x_l - u_l)[(x_l - u_l)^2 + (y_l - v_l)^2], \end{aligned} \tag{7}$$

$$\begin{aligned} &x_r' - u_r = (x_r - u_r) + \\ &k_r(x_r - u_r)[(x_r - u_r)^2 + (y_r - v_r)^2], \end{aligned} \tag{8}$$

where x_l' and x_r' are the actual abscissas in left and right images containing distortions, k_l and k_r are the first order distortion coefficients of the left and right camera lenses.

Using Equations (2), (3), (7) and (8), we obtained:

$$\begin{aligned} &\Delta d = (x_l' - x_r') - (x_l - x_r) = \\ &k_l(x_l - u_l)[(x_l - u_l)^2 + (y_l - v_l)^2] - \\ &k_r(x_r - u_r)[x_r - u_r)^2 + (y_r - v_r)^2]. \end{aligned} \tag{9}$$

We assumed that the lens parameters of the left and right cameras are similar and neglect the centroid ordinate difference of corresponding blocks in the left and right images:

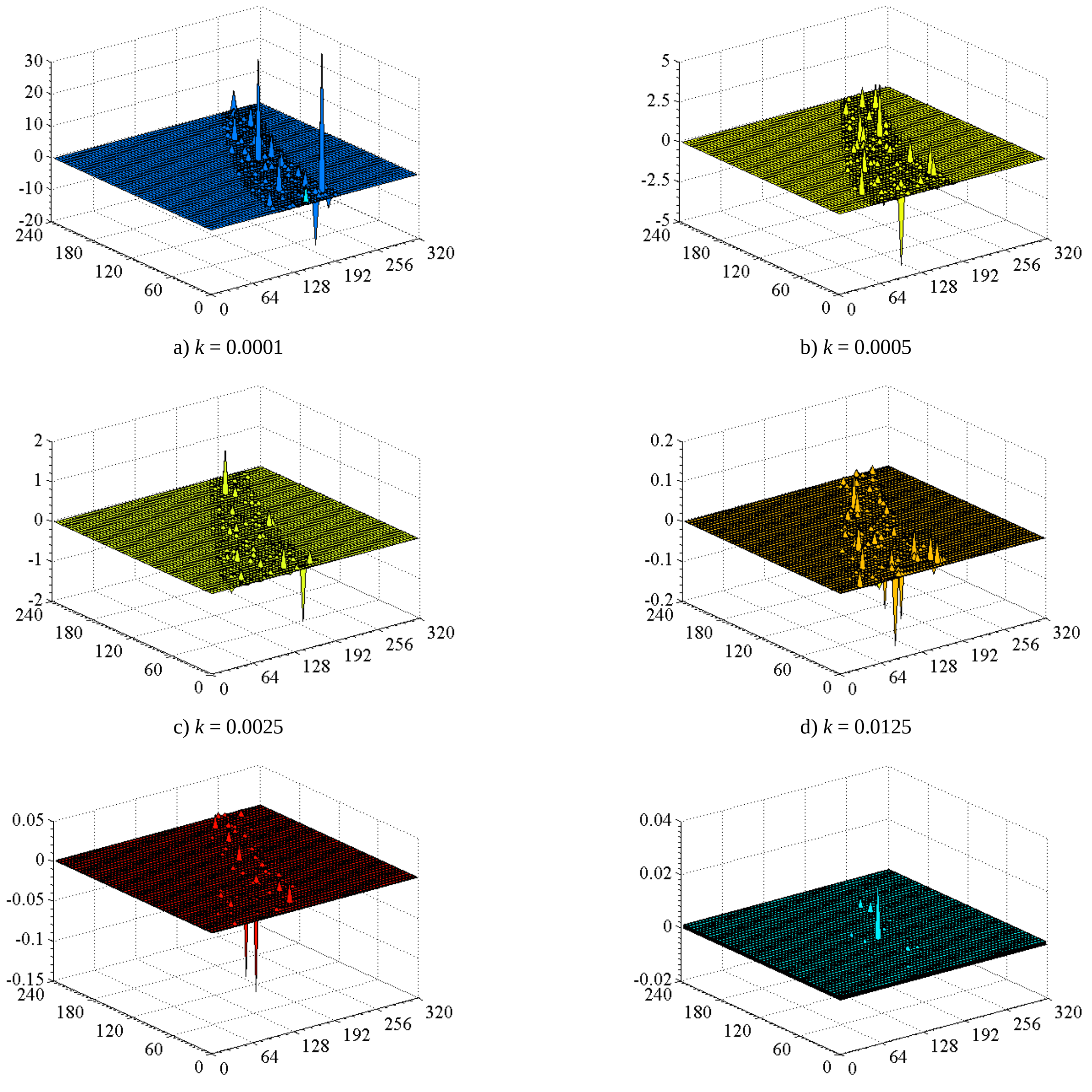
$$u_l \approx u_r = u,\ v_l \approx v_r = v,\ y_l \approx y_r = y,\ v_l \approx v_r = v. \tag{10}$$

Combining Equations (2) and (10), Equation (9) can be written as follows:

$$\begin{aligned} &\Delta d = k(x_l - u)[(x_l - u)^2 + (y - v)^2] - \\ &k(x_r - u)[(x_r - u)^2 + (y - v)^2] = \\ &k(d^3 + 3d^2 - d(x_r - u)^2 + d(y - v)^2). \end{aligned} \tag{11}$$

Finally, we have:

$$\begin{aligned} &g(\Delta d, d) = g(d, x_r, y) \frac{\lambda}{B \cdot f} \cdot \frac{\Delta d}{1 + \Delta d / d} = \\ &\frac{\lambda}{B \cdot f} \cdot \frac{k(d^3 + 3d^2 - d(x_r - u)^2 + d(y - v)^2)}{1 + k(d^2 + 3d - (x_r - u)^2 + (y - v)^2)}. \end{aligned} \tag{12}$$



a) $k = 0.0001$ b) $k = 0.0005$ c) $k = 0.0025$ d) $k = 0.0125$ e) $k = 0.0625$ f) $k = 0.3125$

FIGURE 2 Value range of $f(d, x_r, y)$

3 Numerical simulation and experiment

To observe the range of $g(d, x_r, y)$, image resolution was set as 640×480. Denoting $x_{dist} = x_r - u$, $y_{dist} = y_r - v$, then $x_{dist} \in (1, 319)$, $y_{dist} \in (1, 239)$. The value of k which means the first order radial distortion coefficient is less than 1, so we took a candidate value set as {0.0005, 0.0025, 0.0125, 0.0625}. The camera system we adopted was STH-DCSG-VAR from Videre Design, parameters were obtained as $f = 4.6\,mm$, $\lambda = 0.006\,mm$, $B = 135\,mm$.

By setting the depth estimation extent as 600mm to 4000mm, the range of d was determined as (25.875, 172.5) based on Equation (1). Then the simulation results were obtained (Figure 2).

Points in every image from Figure 2 represented the overlap of data when d varied from 1 to 319. It can be learned that the values of $g(d, x_r, y)$ kept approximately invariable (flat areas) and the change of k only exerts significant influence on the value of breaking points which appeared sparsely. The depth estimation value surging caused by the breaking points can be suppressed by using Kalman Filter. Therefore, Equation (6) is basically suitable for depth estimation error modelling in engineering applications.

Regular blocks of different shapes and sizes were used as objects to evaluate the proposed error model for depth estimation. Part of the results is given in Table. 1.

TABLE 1 Data from testing experiment (Unit: cm)

Cube				Sphere			
Side length=7		Side length=10		Radius=3.5		Radius=5	
z	z’	z	z’	z	z’	z	z’
530	486	550	416	730	742	545	516
930	929	950	856	1130	1205	945	960
1330	1406	1350	1327	1530	1700	1345	1432
1730	1923	1750	1821	1930	2228	1745	1943
2130	2451	2150	2368	2330	2799	2145	2464
2530	3048	2550	2963	2730	3416	2545	3060
2930	3690	2950	3589	3130	4142	2945	3652
3330	4378	3350	4249	3530	4768	3345	4306

All the testing data were utilized for regression. The quadratic regression curve was shown in Figure 3. The regressed coefficients for model $z' = C_1 z + C_2 z^2$ were $C_1 = 0.997$ $C_1 = 6.8 \times 10^{-5}$. It can be seen that the regression result basically coincided with the proposed error model.

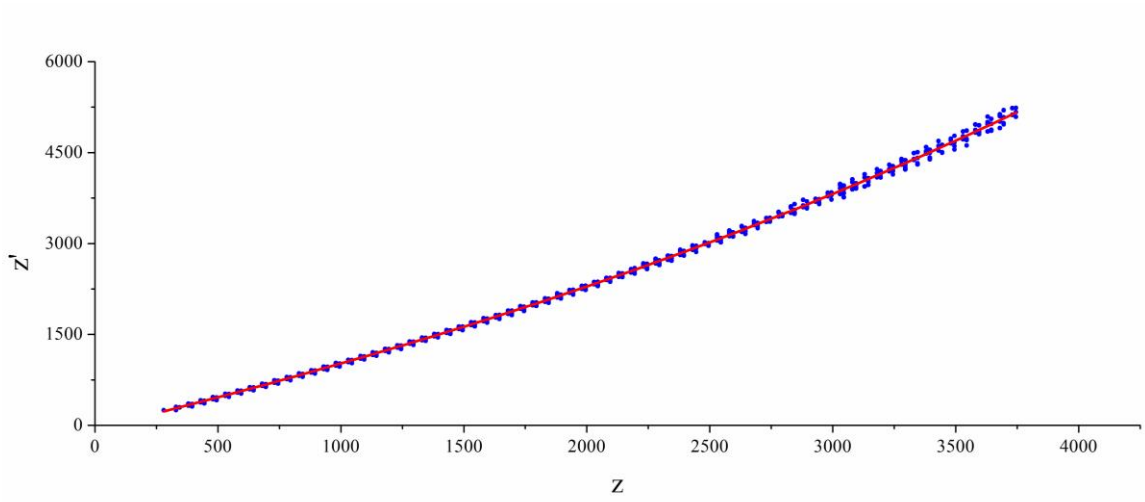


FIGURE 3 All data of validation experiment

4 Conclusions

To facilitate the stereo vision-based obstacle avoidance for mobile robot in poor-textured scenes, we proposed a strategy which needed not the conventional distortion rectification and area-based stereo matching. By image segmentation and obstacle centroid coordinates extraction, abscissa difference was used as parallax to estimate the depth of obstacles. Analytical deduction and numerical simulation were given to prove the feasibility of the hypothesis model. Regression validated the error model in subsequent experiment.

The depth estimation error was larger than the conventional method but still acceptable for mobile robot obstacle avoidance while the processing and calculating time was obviously shortened. The proposed strategy was feasible in general. In future works, the authors will use the proposed strategy combining Kalman Filter for mobile robot indoor obstacle avoidance. Efforts will also be put into improving the proposed model to reduce depth estimation error and processing time.

Acknowledgements

The authors are grateful to the financial support of the Science Fund for Creative Research Groups from the National Natural Science Foundation of China (No. 51221004) and the Program for Zhejiang Leading Team of S&T Innovation (No.2010R50036).

References

[1] Lai X B, Wang HS, Xu Y H 2012 A Real-time Range Finding System with Binocular Stereo Vision *International Journal of Advanced Robotic Systems* **9**(27) 1-9

[2] Wang T, Zheng N N, Xin J M, Ma Z 2011 Integrating Millimeter Wave Radar with a Monocular Vision Sensor for On-road Obstacle Detection Applications *Sensors* **11**(9) 8992-9008

[3] Stelzer A, Hirschmuller H, Gorner M 2012 Stereo-vision-based Navigation of a Six-legged Walking Robot in Unknown Rough Terrain *International Journal of Robotics Research* **31**(4) 381-402

[4] Antunes M, Barreto J P, Premebida C, Nunes U 2012 Can Stereo Vision Replace a Laser Rangefinder? *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* Vilamoura Portugal Oct 2012 5183-90

[5] Donate A, Liu X, Collins E G 2012 *IEEE Transactions on System, Man, and Cybernetics-Part B: Cybernetics* **41**(1) 183-95

[6] Verma R, Singh H S, Verma A K 2013 Depth Estimation from Stereo Images Based on Adaptive Weight and Segmentation *Journal of The Institution of Engineers* **93**(4) 223-9

[7] Fickel G P, Jung C R, Malzbender T, Samadani R, Culbertson B 2013 *IEEE Transaction on Image Processing* **22**(9) 3353-65

[8] Brown M, Hua G, Winder S 2011 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33(1) 2011 43-57

[9] Zou XJ, Zou HX, Lu J 2012 Virtual Manipulator-based Binocular Stereo Vision Positioning System and Errors Modelling *Machine Vision and Applications* 23(1) 43-63

[10] Rodriguez J J, Aggarwal J K 1990 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12(5) 467-70

[11] Fooladgar F F, Samavi S, Soroushmehr S M R, Soroushmehr R 2013 *IEEE Sensors Journal* **13**(11) 4236-46

[12] Llorca D F, Sotelo M A, Parra I 2010 Error Analysis in A Stereo Vision-Based Pedestrian Detection Sensor for Collision Avoidance Applications *Sensors* **10**(4) 3741-58

[13] Chen H D, Chang K C, Agate C S 2013 *IEEE Transactions on Aerospace and Electronic Systems* **49**(2) 840-56

[14] Xiong G M, Li X, Xi J Q 2010 Object Distance Estimation Based on Stereo Vision and Color Segmentation with Region Matching *Advances in Visual Computing* **6455** 368-76

[15] Wang J H, Shi F H, Zhang J 2008 A New Calibration Model of Camera Lens Distortion *Pattern Recognition* 41(2) 607-15

Authors

Bo Jin

Current position, grades: Doctor of Mechanical Engineering, Professor, Zhejiang University.
University studies: B. E. and Ph. D on Fluid Power Transmission and Control awarded by Zhejiang University respectively in 1993 and 1998.
Research interest: deep-sea mechatronic equipment, electric hydraulic control system, intelligent robot, embedded system and applications.
Publications: Over 70 articles, 16 patents and software copyrights, 1 project awarded National Technical invention second prize and 2 projects awarded Provincial S&T progress first prize (Principle person)

Lijun Zhao

Current position, grades: Ph. D. candidate on Mechatronic Engineering, Zhejiang University.
University studies: B. E. (Mechanical Engineering), Lanzhou Jiaotong University (2003). M.E. (Agriculture Mechanization Engineering), Jilin University (2008).
Research Interests: Machine vision and application, mobile robot navigation and control, industrial automation and nondestructive detection.

Shiqiang Zhu

Current position, grades: Doctor of Mechanical Engineering, Professor, Vice-Director of Zhejiang University Robotics Research Center, Vice-President of Zhejiang Ocean University.
University studies: B. E., Fluid Power Transmission and Control, 1988, Zhejiang University; M. E., Mechatronic Engineering, 1991, Beijing Institute of Technology; Ph. D., Mechatronic Engineering, 1994, Zhejiang University.
Research interest: robotics, electro-hydraulic automatic control system, semi-physical simulation technology.
Publications: over 80 articles, 14 patents.