# Application of Active Self-Landmarking to Camera Calibration

Yanfei Liu and Carlos Pomalaza-Ráez

***Abstract*—Camera calibration in an outdoor environment often relies on natural landmarks, and algorithms that are computational expensive. This paper presents a fast camera calibration method that uses robots themselves as landmarks to each other. The active landmark made of LEDs, mounted on the robots, can be turned on/off through available wireless link. Experimental measurements and validation were conducted using a low-resolution, low-computation power smart camera, the CMUcam3 camera. The experimental results show that the estimates of the intrinsic camera parameters converge to the correct values as the robot moves and more measurements are used.**

***Index Terms*—Computer vision, robot sensing, data fusion, sensory networks.**

## I. Introduction

The paradigm used in this paper is similar to the one employed in [1] where a group of mobile robots equipped with sensors measure their positions relative to one another. In [1] the problem being addressed is self-localization whereas our objective here is to have a fast camera calibration procedure. Camera calibration in a mobile robotic network in an unknown environment often relies on landmarks. Landmarks can be natural or artificial. Natural landmarks are selected from some salient regions in the scene. The processing of natural landmarks is usually a difficult computational task. Artificial landmarks are manmade, fixed at certain locations, and of certain pattern, such as circular [2], [3], patterns with barcodes [4], or color pattern with symmetric and repetitive arrangement of color patches [5]. Compared with natural landmarks, artificial landmarks usually are simpler; provide a more reliable performance; and work great for the indoor environment. However it's not an option for most outdoor applications due to the complexity and expansiveness of the fields that robots traverse. Since the robots travel together, one possible solution is to use each other as landmarks.

Using passive landmarks with invariant features such as circular shapes [3] or with simple patterns [4] that are quickly recognizable under a variety of viewing conditions is one option. Whereas these methods have provided good results within indoor scenarios their application to unstructured outdoor environments is complicated by the presence of time varying illumination conditions as well as dynamic objects present in the images. To overcome these drawbacks we propose the use of active landmarks. The current state of LED technology allows for low-power and relative high luminance. Depending on the constraints imposed by the robot's shape and dimensions one or more LEDs can be located on its outer surface. Through wireless links the robots can schedule when the LEDs can be turned on and off to match the periods when the cameras are capturing images for image differencing.

Camera calibration is an important step in computer vision applications, in particular when it is wanted to extract metric information from the 2D images. There is a large body of work on camera calibration techniques developed by the photogrammetry community as well as by computer vision researchers. Most of the techniques assume that the calibration process takes place on a very structured environment, i.e. laboratory setup, and rely on well defined 2D or 3D calibration objects. The use of 1D objects as well as self calibration techniques usually comes with the price of an increase in the computation complexity.

Our work is along the lines of the one described in [6] where closed-form solutions are developed for a calibration method that uses a 1D object. In [6] numerous observations of a 1D object are used to compute the camera calibration parameters. The 1D object is a set of 3 collinear well defined points. The distances between the points are known. The observations are taken while one of the end points remains fixed as the object moves. Whereas the method proposed in [6] is proven to work well in a structured scenario it has several disadvantages if intended to be used in an unstructured outdoors scenario. Depending on the nature of the outdoor scenario, e.g. planetary exploration, having a moving 1D object of a particular length might not be cost effective or even feasible. Our method makes use of a network of robots that can communicate with each other. Thus, it can be implemented in a variety of outdoor environments.

The remainder of the paper is organized as follows. Section II describes the proposed mathematical camera calibration model. Section III presents the validation experiments and the results. Finally the discussion and conclusions are given in Section IV.

Manuscript received on June 25, 2012; revised August 10, 2012. This work was supported by the Indiana Space Consortium through Early Career Faculty and Research Initiation Program.

The authors are with the Department of Engineering, Indiana University – Purdue University, Fort Wayne, IN 46809 USA (e-mail: liu@engr.ipfw.edu, raez@ipfw.edu).

## II. Camera Calibration Model

### A. Notation

For the pinhole camera model (Fig. 1) a 2D point is denoted as $\boldsymbol{a}_i = [a_{ix} \quad a_{iy}]^T$. A 3D point is denoted as $\boldsymbol{A}_i = [A_{ix} \quad A_{iy} \quad A_{iz}]^T$. In Fig. 1 $\boldsymbol{p} = [p_y \quad p_y]^T$ is the

point where the principal axis intersects the image plane. Note that the origin of the image coordinate system is in the corner. $f$ is the focal length.

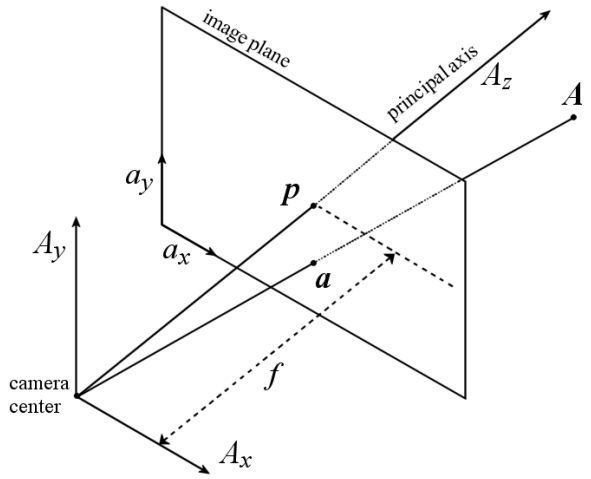



Fig. 1. Normalized camera coordinate system

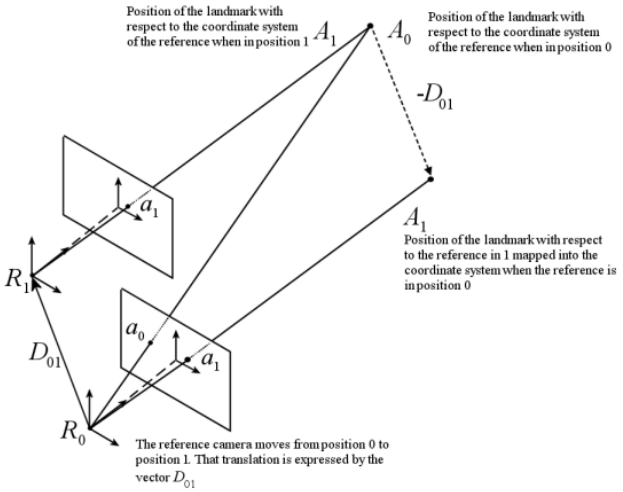



Fig. 2. Changes in the image coordinated when the reference camera or the landmark moves

The augmented vector $\tilde{\boldsymbol{a}}_i$ is defined as $\tilde{\boldsymbol{a}}_i = [a_{ix} \quad a_{iy} \quad 1]^T$. In the same manner $\tilde{\boldsymbol{A}}_i$ is defined as $\tilde{\boldsymbol{A}}_i = [A_{ix} \quad A_{iy} \quad A_{iz} \quad 1]^T$. The relationship between the 3D point $\boldsymbol{A}_i$ and its projection $\boldsymbol{a}_i$ is given by,

$$z_{\boldsymbol{A}_i}\tilde{\boldsymbol{a}}_i = \boldsymbol{K}[\boldsymbol{R} \quad \boldsymbol{t}]\tilde{\boldsymbol{A}}_i \tag{1}$$

where $\boldsymbol{K}$ stands for the camera intrinsic matrix,

$$\boldsymbol{K} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

and

$$\boldsymbol{K}^{-1} = \begin{bmatrix} \frac{1}{\alpha} & -\frac{\gamma}{\alpha\beta} & \frac{\gamma v_0 - u_0\beta}{\alpha\beta} \\ 0 & \frac{1}{\beta} & -\frac{v_0}{\beta} \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$[u_0 \quad v_0]$ are the coordinates of the principal point in pixels, $\alpha$ and $\beta$ are the scale factors for the image $u$ and $v$ axes, and $\gamma$ stands for the skew of the two image axes. $[\boldsymbol{R} \quad \boldsymbol{t}]$ stands for the extrinsic parameters and is the rotation and translation that relates the world coordinate system to the camera coordinate system. Without loss of generality we can assume for our analysis that $\boldsymbol{R} = \boldsymbol{I}$ and $\boldsymbol{t} = \boldsymbol{0}$.

If $\gamma = 0$ as it is the case for CCD and CMOS cameras then,

$$\boldsymbol{K} = \begin{bmatrix} \alpha & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

and

$$\boldsymbol{K}^{-1} = \begin{bmatrix} \frac{1}{\alpha} & 0 & -\frac{u_0}{\alpha} \\ 0 & \frac{1}{\beta} & -\frac{v_0}{\beta} \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

The $\boldsymbol{K}$ matrix can also be written as,

$$\boldsymbol{K} = \begin{bmatrix} m_x f & 0 & m_x p_x \\ 0 & m_y f & m_y p_y \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

where $m_x$, $m_y$ are the number of pixels per meter in horizontal and vertical direction.

### *B. Mathematical Model*

The model described in this section is illustrated in Fig. 2. The reference camera is at position $\boldsymbol{R}_i$ while the landmark is located at position $\boldsymbol{A}_i$. The projection of the landmark in the image plane of the reference camera changes when the camera moves from position 0 to position 1 as illustrated in Fig. 2. This motion is represented by the vector $\boldsymbol{D}_{01}$. If instead the landmark moves according to $-\boldsymbol{D}_{01}$, as shown in Fig. 2. and the reference camera does not move, then both the location of the landmark, $\boldsymbol{A}_1$, and its projection on the image, $\boldsymbol{a}_1$, would be the same as in the case when the reference camera moves.

For any location of the landmark, $\boldsymbol{A}_i$, and its projection on the image, $\boldsymbol{a}_i$, If $\tilde{a}_i = \begin{bmatrix} a_{ix} a_{iy} & 1 \end{bmatrix}^T$ with $\boldsymbol{R} = \boldsymbol{I}$ and $\boldsymbol{t} = \boldsymbol{0}$ from eq. (1), then

$$A_i = z_{A_i} K^{-1} \tilde{a}_i \tag{7}$$

also define

$$D_{ij} = A_j - A_i = \begin{bmatrix} d_{ijx} d_{ijy} d_{ijz} \end{bmatrix}^T \tag{8}$$

The magnitudes of $A_i, A_j,$ and $D_{ij}$ ( $L_{A_i}, L_{A_j},$ and $L_{D_{ij}}$, respectively) can be estimated using the strength of the received signal. Also for $D_{ij}$ it is possible to estimate $L_{D_{ij}}$ using the data from the robot navigational systems. Both estimation methods, signal strength on a wireless link and navigational system, have certain amount of error that should be taken into account in the overall estimation process.

$$\begin{aligned} L_{A_j}^2 = A_j{}^T A_j = \left(A_i{}^T + D_{ij}{}^T\right)\left(A_i + D_{ij}\right) = \\ = A_i{}^T A_i + A_i{}^T D_{ij} + D_{ij}{}^T A_i + D_{ij}{}^T D_{ij} \end{aligned} \tag{9}$$

$$L_{A_j}^2 = L_{A_i}^2 + L_{D_{ij}}^2 + 2{D_{ij}}^T A_i \tag{10}$$

$$\begin{aligned} {D_{ij}}^T A_i &= \frac{L_{A_j}^2 - L_{A_i}^2 - L_{D_{ij}}^2}{2} \\ &= {D_{ij}}^T z_{A_i} \begin{bmatrix} \frac{1}{\alpha} & 0 & -\frac{u_0}{\alpha} \\ 0 & \frac{1}{\beta} & -\frac{v_0}{\beta} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{ix} \\ a_{iy} \\ 1 \end{bmatrix} \end{aligned} \tag{11}$$

Define $\delta_{ij}$ as,

$$\begin{aligned} \delta_{ij} &= \frac{L_{A_j}^2 - L_{A_i}^2 - L_{D_{ij}}^2}{2} = \\ &= z_{A_i} \begin{bmatrix} d_{ijx} & d_{ijy} & d_{ijz} \end{bmatrix} \begin{bmatrix} M_1 & 0 & M_2 \\ 0 & M_3 & M_4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{ix} \\ a_{iy} \\ 1 \end{bmatrix} \end{aligned} \tag{12}$$

where,

$$M_1 = \frac{1}{\alpha} M_2 = -\frac{u_0}{\alpha} M_3 = \frac{1}{\beta} M_4 = -\frac{v_0}{\beta} \tag{13}$$

for $0 \le i < j \le N$

where $N$ is the number of locations where the landmark moves to

$$\begin{aligned} \delta_{ij} = z_{A_i} \big[ & a_{ix} d_{ijx} M_1 + d_{ijx} M_2 + \\ & a_{iy} d_{ijy} M_3 + d_{ijy} M_4 + d_{ijz} \big] \end{aligned} \tag{14}$$

$z_{A_i}$ is the projection of the vector $A_i$ on the $z$-axis. The value of $z_{A_0}$ itself can be using the navigation system as the robot takes the first measurement position. Then $z_{A_i}$ can be calculated using the mobile robot's navigation system by adding the movement along $z$-axis from one location to another. Assuming then that $z_{A_i}$ has been estimated,

$$\frac{\delta_{ij}}{z_{A_i}} - d_{ijz} = \begin{bmatrix} a_{ix} d_{ijx} & d_{ijx} & a_{iy} d_{ijy} & d_{ijy} \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix} \tag{15}$$

Let's define $\lambda_{ij}$ as,

$$\lambda_{ij} = \frac{\delta_{ij}}{z_{A_i}} - d_{ijz} \tag{16}$$

$$\text{for } 0 \le i < j \le N$$

Then,

$$\lambda_{ij} = \boldsymbol{c}_{ij}^T \boldsymbol{x} \tag{17}$$

where $\boldsymbol{c}_{ij} = \begin{bmatrix} a_{ix} & d_{ijx} & a_{iy} & d_{ijy} & d_{ijy} \end{bmatrix}^T$

$$\text{and } \boldsymbol{x} = [M_1 M_2 M_3 M_4]^T$$

If the landmark moves to $N$ locations, $A_1, A_2, \cdots, A_N$, the corresponding equations can be written as,

$$\boldsymbol{\Lambda} = \mathbf{C}\boldsymbol{x} \tag{18}$$

where $\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_{12} \lambda_{13} \cdots \lambda_{1N} \lambda_{23} \cdots \lambda_{(N-1)N} \end{bmatrix}^T$

and $\mathbf{C} = \begin{bmatrix} \boldsymbol{c}_{12} \boldsymbol{c}_{13} \cdots \boldsymbol{c}_{1N} \boldsymbol{c}_{23} \cdots \boldsymbol{c}_{(N-1)N} \end{bmatrix}^T$

The $N$ locations are cross-listed to generate a number of $N(N-1)/2$ points in the equations as shown in Fig. 3. The least-squares solution for $\boldsymbol{x}$is,

$$\boldsymbol{x} = [\mathbf{C}^T \mathbf{C}]^{-1} \mathbf{C}^T \boldsymbol{\Delta} \tag{19}$$

Once $\boldsymbol{x}$ is estimated the camera intrinsic parameters can be easily computed.

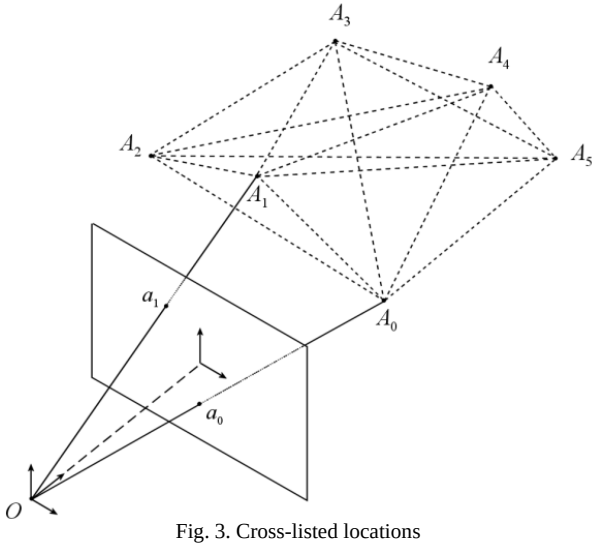



Fig. 3. Cross-listed locations

## III. Experimental Validation

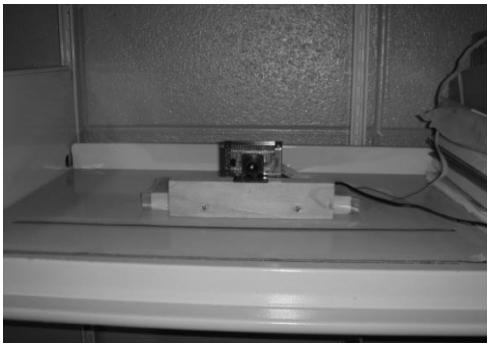
To test the calibration method itself, we used common construction tools, such as tape measures, rulers, plumb-blob, to carefully measure the coordinates of each landmark location in the camera coordinate system and distances between the reference camera and the landmark. Using a laser range finder we estimated that the errors incurred using those construction tools are in the order of $\pm$ 2 or 3 cm. The same range of estimate error is incurred when using wireless communication technology such as Ultra-wideband (UWB).

When this calibration method is implemented in a mobile robotic network, the actual coordinates of each location cannot be known. In our mathematical model, the known variables are the vectors between the landmarks' locations in the reference camera coordinate system and the image coordinates of each landmark. In our experiments, the measurements of the landmarks' coordinates are not used directly in the calculation. Instead these measurements are used to calculate the vectors between the landmarks. When this calibration method is used in a mobile robotic network, these vectors are obtained by the robots' navigation system.

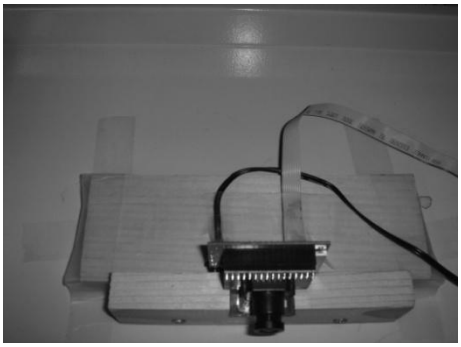
The smart-camera we used in the experiments is a CMUcam3 [7]. The image sensor is an Omnivision OV6620 CMOS camera board that outputs a stream of 8-bit RGB or YCbCr color pixels. The OV6620 supports a maximum resolution of 352x287 at 50 frames per second. The focal length of this camera is 2.8~4.9mm. $m_x$is 1/9.0μm and $m_y$ is

1/8.2μm.

The CMUcam3 was mounted in a regular office environment. A wood frame was built to support the camera in a way that the *z*-axis (principle axis) is in the horizontal direction. Fig. 4. below shows the front and top view of the CMUcam3 and the mounting structure.



(a) Front view



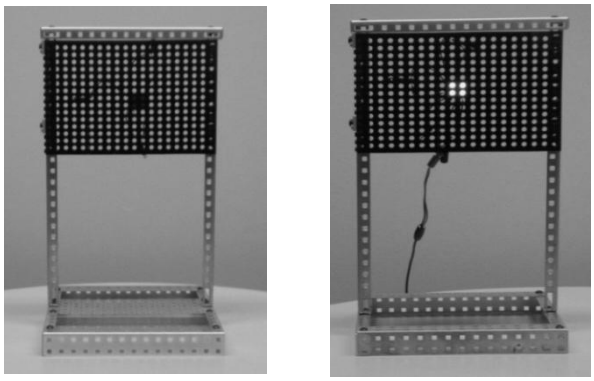
(b) Top view

Fig. 4. The CMUcam3

The active landmark was formed using the metal structure parts from the VEX robotics design system [8] and LEGO bricks with holes. With the possible wireless communication capabilities, the robots can turn on and off the LEDs whenever needed to form visible landmarks. Fig. 5 shows the pictures of the robot frame where the LEDs are in ON and OFF state.

The metal frame with the active landmark was placed in different locations in the room. For our experiment twelve locations were chosen so that the landmarks were spread out in the image plane. A newly developed efficient blob finding algorithm [9] was used to automatically find the landmark anywhere in a scene and then calculate the centroid of the landmark. Fig. 6. shows the pictures of one of the landmark locations with the LEDs turned on/off and the output from the blob finding algorithm.

The measurements of the landmarks in the twelve locations are shown in Table I. $A_x$, $A_y$, and $A_z$ are the coordinates in the camera coordinate system. L is the magnitude of the *A* vector. $a_x$ and $a_y$ are the image coordinates. To make most use of measurements taken $n$ locations they can be used to generate a number of $n(n-1)/2$ equations as shown in Fig. 3. Thus the twelve points can be used to generate a maximum of (12 x 11)/2=66 equations. In order to compare the results of the calibration model using different numbers of measurements, our calculation used 5~12 locations that generates 10~66 equations. The calculation results are shown in Table II. In the datasheet of the CMUcam3, a range was given for the focal length $f$ (2.8~4.9mm). With the value of $m_x$ and $m_y$, we can calculate the range of α (311.1~544.4) and β (341.5~597.6). It is difficult to know what the exact value of $f$ is, thus the exact values of α and β cannot be known either. However, the ratio of α/β is known and is equal to $m_x/m_y$(0.91). The relative errors of the estimation of the intrinsic parameters are shown in Fig. 7. The estimation results show that the estimates of the parameters converge to the correct values as more measurements are used.

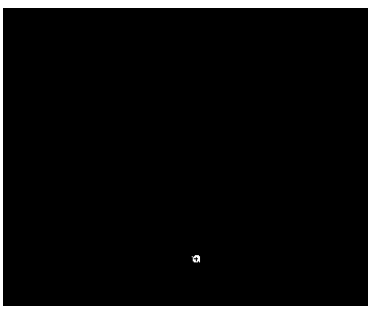


(a) LEDs OFF (b) LEDs ON

Fig. 5. Active landmarks

(a) Landmark OFF

(b) Landmark ON



(c) Centroid of the landmark

Fig. 6. Pre- and post-processed imaged by the CMUcam3

TABLE I: MEASUREMENT

| | Ax (cm) | Ay (cm) | Az (cm) | L (cm) | ax(pixels) | ay(pixels) |
|---|---|---|---|---|---|---|
| **1** | -3.4 | -41.0 | 184.5 | 188.0 | 166 | 45 |
| **2** | -61.0 | -41.0 | 189.0 | 199.0 | 47 | 49 |
| **3** | -21.0 | -41.0 | 230.0 | 232.0 | 145 | 63 |
| **4** | -66.0 | -41.0 | 229.0 | 239.0 | 61 | 65 |
| **5** | -10.0 | 2.0 | 176.6 | 176.6 | 153 | 149 |
| **6** | -74.0 | 2.0 | 177.8 | 189.6 | 11 | 147 |
| **7** | -12.0 | 2.0 | 224.0 | 224.0 | 151 | 148 |
| **8** | -67.0 | 2.0 | 228.4 | 235.2 | 56 | 147 |
| **9** | 33.0 | -46.0 | 264.4 | 271.5 | 228 | 70 |
| **10** | 29.0 | -3.0 | 287.0 | 287.7 | 218 | 143 |
| **11** | 50.0 | 31.5 | 272.5 | 281.3 | 251 | 197 |
| **12** | -66.0 | 36.5 | 223.5 | 232.7 | 60 | 215 |

TABLE II: CALIBRATION RESULTS

| # of Data points | Image sets used | α | β | $u_0$ | $v_0$ | α/β |
|---|---|---|---|---|---|---|
| 10 | 1→5 | 391.9 | 440.0 | 174.6 | 143.0 | 0.89 |
| 15 | 1→6 | 392.6 | 368.1 | 175.4 | 128.3 | 1.07 |
| 21 | 1→7 | 394.0 | 386.1 | 175.5 | 132.0 | 1.02 |
| 28 | 1→8 | 393.3 | 371.2 | 175.3 | 130.0 | 1.06 |
| 36 | 1→9 | 395.0 | 452.6 | 175.2 | 145.3 | 0.87 |
| 45 | 1→10 | 396.0 | 451.0 | 175.0 | 145.4 | 0.88 |
| 55 | 1→11 | 397.5 | 442.0 | 175.3 | 144.4 | 0.90 |
| 66 | 1→12 | 399.7 | 442.5 | 176.0 | 144.0 | 0.90 |
| **Intrinsic parameters** | | | | **176** | **143** | **0.91** |

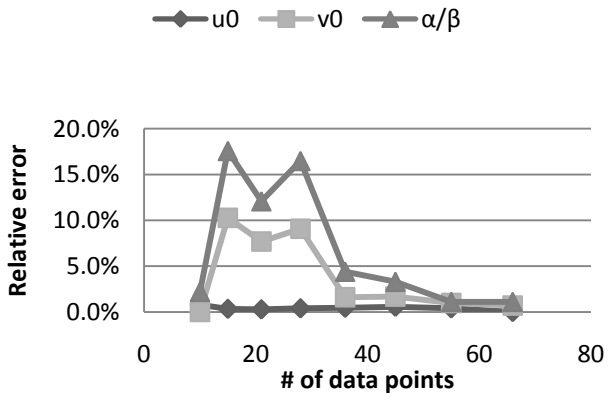



Fig. 7. Relative error of the intrinsic parameters

## IV. CONCLUSION

In this paper, a new method for fast camera calibration in outdoor environments is presented and tested using a smart-camera, the CMUcam3 camera. This method can be implemented in a camera-equipped wireless mobile robotic network, where the robots themselves are used as each other's landmarks. The distances between the robots can be estimated using wireless communication links supported by standard protocols. Active landmarks made of LEDs, which can be turned on and off through wireless communications, is proposed to realize fast calibration computation.

## ACKNOWLEDGMENT

The author would like to thank Indiana Space Consortium for their support to the project.

## REFERENCES

[1] R. Kurazume, S. Hirose, S. Nagata, and N. Sahida, "Study on Cooperative Positioning System: Basic Principle and Measurement Experiment," presented at the 1996 IEEE International Conference on Robotics and Automation, Minneapolis, Minnesota, April 1996.

[2] C. Lin and R. L. Tummala, "Mobile Robot Navigation Using Artificial Landmarks," *Journal of Robotics Systems*, vol. 14, no. 9, pp. 93-106, 1997.

[3] B. Zitova and J. Flusser, "Landmark recognition using invariant features," *Pattern Recognition Letters*, vol. 20, pp. 541-547, 1999.

[4] A. Briggs, D. Scharstein, D. Braziunas, C, Dima, and P. Wall, "Mobile Robot Navigation Using Self-Similar Landmarks," in *Proceeding of the 2000 IEEE International Conference on Robotics and Automation*, April 2000, pp. 1428-1434.

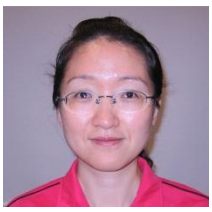
[5] K. J. Yoon and I. S. Kweon, "Landmark design and real-time landmark tracking for mobile robot localization," in *Proceedings of The International Society for Optical Engineering*, Boston, USA, vol. 4573, no. 21, October 29-November 2, 2001, pp. 219-226.

[6] Z. Zhang, "Camera Calibration With One-Dimensional Objects," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 7, pp. 892-899, 2004.

[7] CMUcam3. [Online]. Available: http://www.cmucam.org.

[8] VEX Robotics Design Systems. (September 2010). [Online]. Available: http://www.vexrobotics.com.

[9] Y. Liu and C. A. Pomalaza-Ráez, "On-Chip Body Posture Detection for Medical Care Applications Using Low-Cost CMOS Cameras," *Journal of Integrated Computer-Aided Engineering*, vol. 17, no. 1, pp. 3-13, 2010.



**Yanfei Liu** received a BS (1996) in Electrical Engineering from Shandong Institute of Architecture and Engineering, Jinan, China, a MS (1999) in Electrical Engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, and a PhD (2004) in Electrical Engineering from Clemson University in South Carolina. In 2005 she joined the Department of Engineering at Indiana University – Purdue University Fort Wayne (IPFW), where she currently is an associate professor of electrical and computer engineering. She is the author/co-author of more than 20 publications in technical journals and conferences. Her Ph.D. work was on vision guided dynamic manipulation for industrial robots. Her current research focuses on autonomous mobile robots and distributed sensing. She is also particularly interested in distributed sensing uthrough low-cost, low-power consumption smart cameras with wireless capabilities. Professor Liu is a member of the IEEE. She is also the recipient of the IEEE Real World Engineering Award in 2010, for the freshmen engineering project "Energy Scavenging through Vibrations."

**Carlos Pomalaza-Ráez** received a BSME and a BSEE degree from Universidad Nacional de Ingeniería, Lima, Perú, in 1974, and the M.S. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1977 and 1980, respectively. He is an electrical engineering professor at Indiana University - Purdue University Fort Wayne, Indiana, USA and a visiting professor at the University of Oulu, Finland.He has been a faculty member of the University of Limerick, Ireland, and of Clarkson University, Potsdam, New York. He has also been a member of the technical staff at the Jet Propulsion Laboratory of the California Institute of Technology.
Professor Pomalaza-Ráez is a senior member of the IEEE. In 2003 and 2004, under the auspices of a Nokia-Fulbright Scholar Award, he was a visiting professor at the Centre for Wireless Communications, University of Oulu, Finland.