

MODEL-BASED CALIBRATION FOR SENSOR NETWORKS

J. Feng, S. Megerian, M. Potkonjak

Department of Computer Science
University of California, Los Angeles
Los Angeles, California 90095, USA

ABSTRACT

Calibration is the process of mapping raw sensor readings into corrected values by identifying and correcting systematic bias. Calibration is important from both off-line and on-line perspectives. Major objectives of calibration procedure include accuracy, resiliency against random errors, ability to be applied in various scenarios, and to address a variety of error models. In addition, a compact mapping function is attractive in terms of both storage and robustness. We start by introducing the non-parametric statistical approach for conducting off-line calibration. After that, we present the non-parametric statistical percentile method for establishing the confidence interval for a particular mapping function.

Furthermore, we propose the first model-based on-line procedure for calibration. The calibration problem is formulated as an instance of nonlinear function minimization and solved using the standard conjugate gradient approach. A number of trade-offs between the effectiveness of calibration and noise level, latency, size of network and the complexity of phenomena are analyzed in a quantitative way. As a demonstration example, we use a system consisting of photovoltaic optical sensors.

1. INTRODUCTION

Sensor calibration is an inevitable requirement due to the natural process of device decadence and imperfection. It is particularly important in wireless distributed sensor networks, because it is difficult to conduct manual calibration. Calibration is a technically challenging task mainly due to the existence of random noise and the absence of suitable error models. In addition, in many applications there are constraints in terms of latency. The direct ramification is that in these situations the calibration procedure has to be conducted rapidly within a certain time frame.

Our goal is to study calibration within the framework of distributed sensor systems. The main technical objectives include proposing several calibration error models, collecting and analyzing real-life sensor data required for calibration, identifying relevant metrics for high quality calibration, developing off-line and on-line calibration procedures, providing sound statistical evaluation mechanisms to measure how well the calibration procedures work, and conducting comprehensive calibration simulations and experiments using optical sensors.

The rest of the paper is organized in the following way. We survey the state-of-the-art calibration techniques and some of the most recent developments in related fields in Section 2. After that, we discuss the point-lights equation-based. In Section 4, we introduce the off-line calibration technique, and apply the point-lights model as an example to illustrate the procedure. The on-line

calibration technique is presented in Section 5, using the point-lights model as an a demonstration case study.

2. RELATED WORK

In this Section, we briefly survey the related work along the following lines: sensor data calibration, statistical modeling, nonlinear programming and function minimization.

Wireless sensor networks have the potential of serving as the bridge between the current Internet and the physical world. In order to have accurate and reliable data, effective mechanisms and techniques for calibration are of prime importance. The existing state-of-the-art calibration techniques for sensor networks include [22][20][21][12]. Whitehouse and Culler [22] describe why traditional calibration methods are not suitable for ad-hoc distributed and localized systems such as sensor networks, and propose formulation of a calibration technique for localization as a parameter estimation problem. Bychkovskiy, Megerian, Estrin and Potkonjak [20] introduce a two-phase calibration technique that first finds all the pair-wise calibration functions, and then optimizes the global calibration function based on the pair-wise calibration functions found in phase one.

Welch and Bishop [21] not only propose a new mathematical method for tracking a user's position and orientation for interactive computer graphics, but also introduces an autocalibration technique that can be done concurrently while performing tracking. SCAAT isolates the individual measurements, or measurement dimensions in such a way that individual source and sensor imperfections are more easily identified and addressed. SpotON is a flexible alternative to infrastructure-centric location systems. Hightower, Vakili and Borriello [12] address calibration by using a distance dependent RSSI prediction function. Several methods are proposed for data collection, transmitter and receiver variations.

Calibration of sensors can also be conducted from other perspectives. For example, Elson and Estrin [7] calibrate the sensors from time synchronization point of view. All the sensors that receive a broadcast message calibrate among themselves using the stamped receiving time as a reference. Borenstein and Feng [4] introduce systematic near-optimal calibration methods for measuring and reducing odometry errors in mobile robotics. Hoff and Azuma [13] describe an autocalibration method for electronic compasses that compensates compass distortions. This approach takes the advantage of the redundant heading information computed from rate gyroscopes. The calibration table is automatically adjusted to correct the compass output. Faugeras, Luong and Maybank [10] introduce a camera calibration (finding the internal orientation of a camera) method that does not require the camera motion to be known. The camera is calibrated by

simply pointing it at the environment, selecting and tracking points of interest in the image as the camera moves.

A most popular parametric technique for fitting a set of data to a particular compressed form is to use linear least-square fitting. The technique is optimal under the assumption that the error distribution follows the Gaussian distribution. The standard technique for conducting least square fitting is to use the Singular Value Decomposition (SVD) approach. In addition, various techniques are available for determining whether the least-squares model is appropriate of the given set of data [2][5]. The detailed description of these methods can be found in [16]. In addition to parametric techniques, recently non-parametric techniques that do not depend on any underlying assumptions have been gaining popularity [19]. The standard references include [6][8][9][11].

Linear programming and nonlinear programming have been popular optimization mechanisms since the 1940's. Both of these methods involve three entities: variables, objective function, and constraints. The goal is to find a set of assignments to the variables in such a way that the objective function is minimized or maximized, and at the same time have all the constraints remain satisfied. Standard references for linear programming include [18]. For nonlinear programming, some useful sources include [1][3][14]. We solve nonlinear programming problem by weighting all constraints into the objective function and rely on the conjugate gradient-based unconstrained nonlinear function minimization techniques as the means to solve the instance. Specifically, we used the public available software package WNLIB [15].

3. PRELIMINARIES

Before we present the off-line and on-line calibration procedure, in this Section, we state the adopted models, abstractions and assumptions.

From a network topology point of view, we assume a stationary wireless sensor network that is densely deployed. This assumption is particularly crucial in the case of on-line calibration where sensors rely on each other to calibrate themselves. More specifically, a densely deployed network provides sufficient amount of data redundancy, which is the basis of on-line calibration.

Sensing devices can be quite different in their response characteristics to stimuli. One can model the errors in the measurements by classifying them into two major categories: systematic bias and random noise. Bias is the systematic offset in the amplitude of sensor readings from the true value. We assume that all sensors have constant bias, which means that each sensor is always off by a constant amount from the true value. Noise represents the random component in the error. The sources of noise are external events that influence sensor readings, hardware noise, or other unpredictable transient events. We assume the random noise added in our measurements follow the Gaussian distribution with mean 0 and sigma 3 [17].

We use the point-lights model as an example to demonstrate both the off-line and the on-line calibration procedures. All the simulation and experimentations are conducted with respect to the point-lights model. The formulation of point-lights model is stated as follows:

INSTANCE: A set of n light sensors S where for each node $s_i \in S$, $i = 1, \dots, n$, has the position $(x_{s_i}, y_{s_i}, z_{s_i})$; each sensor has intensity measurement I_{s_i} .

QUESTION: Assuming that k point-light sources in the environment is known, for each light l_j , $j = 1, \dots, k$, find the position $(x_{l_j}, y_{l_j}, z_{l_j})$, and the intensity of the source I_{l_j} .

Figure 1 shows the irradiance of a point light source. Equation (3) is used as the light model each sensor incorporates. I_{s_i} denotes the intensity of light measured by the sensor, I_{l_j} is the true intensity the source has. The model of light at each sensor directly follows by rewriting the equations in Figure 1. For every sensor $s_i \in S$, given its position and orientation, we have the light intensity I_{s_i} expressed in equation (4).

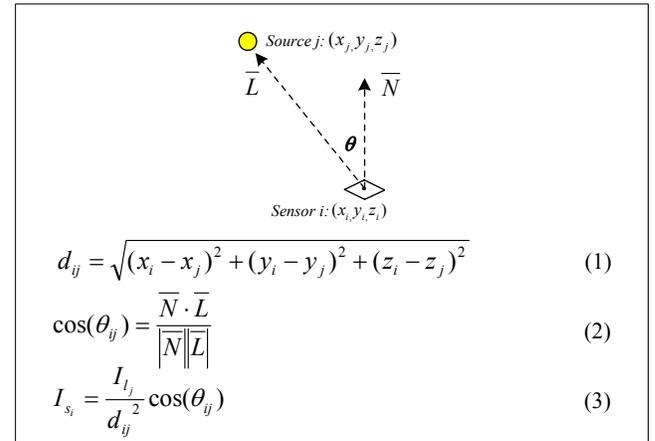


Figure 1. Point-light models and sensor reading

$$I_{s_i} = \sum_{j=1}^k \frac{I_{l_j}}{d_{ij}^2} \cos(\theta_{ij}) \quad (4)$$

In both off-line and on-line experimentations, the data is recorded using a miniature silicon solar cell that converts light impulses directly in to electrical charges (photovoltaic) and a photoconductor. The silicon solar cell generates its own power and does not require any external bias unlike other conventional photo diodes or transistors. More specifically, this silicon cell is mounted on a 0.78cm x 0.58cm x 0.18cm thick plastic carrier and generates about 400mV in moderate light, which is similar to a typical room. An Extech model 407026 commercial digital light meter is also used in the off-line calibration. Detailed description of the light appliances can be found in [23].

4. OFF-LINE CALIBRATION

In this Section, we introduce the off-line calibration technique. We use light sensing as a demonstrative example. After we state the problem formulation and analyze the procedure, we apply the percentile resubstitution method in order to obtain the confidence interval. Finally, the experimental results are presented.

4.1 Problem Formation

In the context of off-line calibration, the collected data has two components: raw sensor readings and data captured by high-quality and high-cost light meters measuring the same set of stimuli at essentially the same positions and the same angles. The second set of data serves as the standards of what the low-cost sensors should measure. The goal of the off-line calibration is to find a compact function that provides the mapping from the raw sensor reading to the correct values. In order to avoid over-fitting and ensure accuracy, the function should simultaneously use a small number of parameters and should have low prediction error. The formal problem definition is stated as follows:

INSTANCE: A set of pairs of real numbers $S = \{ \{r_i, c_i\} \mid r_i \in R = \{r_1, r_2, \dots, r_n\}, c_i \in C = \{c_1, c_2, \dots, c_n\} \}$ where R is the data produced by low-cost sensors; and corresponding C is the data produced by high-accuracy instrument, integer K , and TOL – predefined tolerance constant expressed as a real number are given.

QUESTION: Is there a calibration function $f(R)$ that satisfies the following conditions:

- (i) $|f(r_i) - c_i| \leq \text{TOL}, i=1, \dots, n$
- (ii) $f(R)$ employs no more than K parameters

4.2 Experimental Results

Figure 2 shows a sensor’s pair-wise sensor readings and meter readings over a period of time. The raw sensor readings are shown along the x-axis, which corresponds to set R ; the calibrated high-quality light meter readings are shown along the y-axis, which corresponds to set S .

We have tried a number of alternatives to solve the problem, including fitting using the linear least square technique and employing a number of polynomial and logarithmic functions as parameters. However, the residual discrepancy resulted was too large according to the standard hypothesis validation test. Finally we found that it is beneficial if the data points are separated into two sets, namely $[0 \leq r_i \leq 375]$ and $(375 \leq r_i \leq 1600]$. It was sufficient to fit each interval using a separate linear function. Table 1 shows the two separate calibration functions obtained for the two intervals respectively.

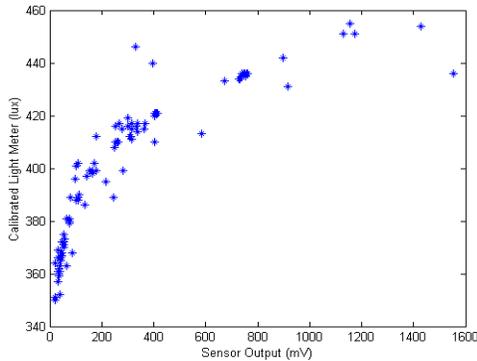


Figure 2. Pair-wise sensor readings and meter readings

Linear fit for data points $r_i = [0, 375]$	$y = 0.184x + 361$
Linear fit for data points $r_i = (375, 1600]$	$y = 0.0281x + 412$

Table 1. Calibration functions for the optical

Figure 3 shows the data points along with the two fitted calibration function stated in Table 1.

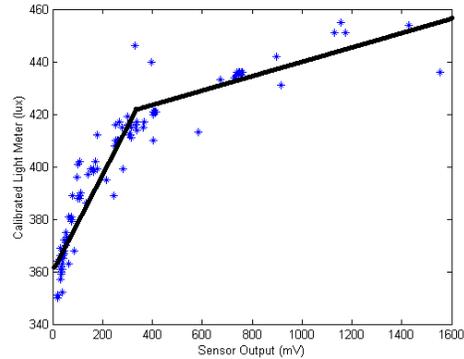


Figure 3. Mapping of sensor readings and meter readings using piece-wise fitted calibration functions

4.3 Solution Evaluation

After calibration functions are established, we applied the non-parametric statistical percentile method to obtain the confidence interval for the calibration functions. More specifically, given a specific R (sensor reading) value, we compare what the calibration function predicts with what C (calibrated light meter) reads, and construct the number of occurrence histogram based on a number of different instances generated. Figure 4 and 5 are the number of occurrence histogram for the two intervals.

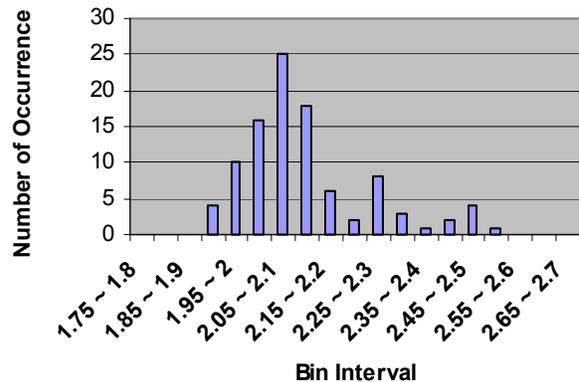


Figure 4. Percentile method histogram for sensor data $r_i = [0, 375]$

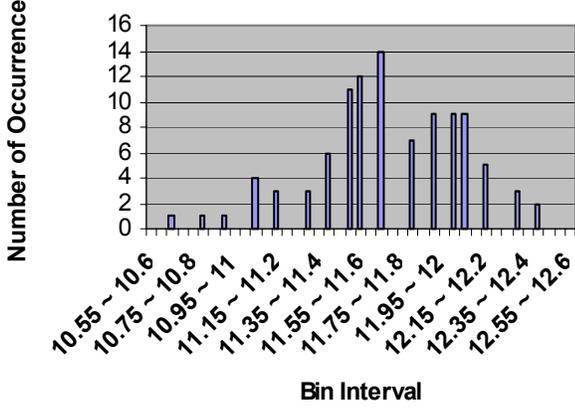


Figure 5. Percentile method histogram for sensor data
 $r_i = (375, 1600]$

5. ON-LINE CALIBRATION

In this Section, we introduce the collaborative on-line calibration technique. While the technique is generic in the sense that it can be applied to a great variety of phenomena and types of sensors, we demonstrate it for the case of point-light sources and light sensors. The technique does not require the presence of accurate light meters to serve as standards as in the case of off-line procedure. We utilize redundancy in sensor deployment in order to enable the calibration process. The technique is also generic with respect to any function of calibration discrepancy. For the sake of simplicity and clarity, we restrict our attention on the case where each sensor reading only has constant bias. We first discuss how the problem can be formulated as an instance of nonlinear function minimization. The section is concluded by presenting both experimental and simulation results as well as their evaluation using percentile method.

5.1 Problem Formation

We first distinguish the notion of measured and calculated intensities. The measured intensity is simply the sensor reading; the calculated intensity is the result obtained applying equation (4). In the perfect situation where there is no error present (both systematic and noise), the readings are identical. However, the presence of error is inevitable. Therefore, the measured intensity is not necessarily equal to the real intensity. Equation (4) can therefore be rewritten as equation (5). $I_{M_{si}}$ denotes the measured intensity, B_i denotes the systematic error that we are trying to correct, and ε_i denotes the random noise. The summation of these three terms equal to $I_{R_{si}}$, which denotes the real intensity.

$$(I_{M_{si}} + B_i + \varepsilon_i) = \sum_{j=1}^k \frac{I_{lj,t}}{d_{ij,t}^2} \cos(\theta_{ij,t}) \quad (5)$$

Furthermore, we rewrite equation (5) in another form:

$$\varepsilon_i = \sum_{j=1}^k \frac{I_{lj,t}}{d_{ij,t}^2} \cos(\theta_{ij,t}) - I_{M_{si}} - B_i \quad (6)$$

This equation contains five unknown: the coordinates of light source j (x_{ij} , y_{ij} , z_{ij}); the intensity of light j I_{lj} ; and finally the bias

associated with sensor i B_i . Since the goal is to minimize the discrepancies, the objective function should be defined as a function of all ε_i .

$$obj: \quad \min f(\varepsilon_i) \quad i = 1, \dots, n \quad (7)$$

Once the objective function is specified (we used various vector norms for this purpose), we rely on the conjugate gradient-based nonlinear function minimizer [15] to solve the problem instance.

Note that for a given set of sensor measurements at time t , this formulation results in an under-determined set of equations since any solution to the model can trivially be satisfied by setting all ε parameters equal to the discrepancy between the physical sensor measurement and the corresponding value enforced by the phenomenon model (equation (4)). In order to adequately constrain this problem, we introduce the concept of *time steps*, (e.g. consider measurements from different times t_α and t_β). More specifically, the objective function is now minimizing ε_i 's of all sensors at multiple time instances instead of just one time instance as in the case of equation (6). The formulation can be rewritten as follows where the subscript t indicates a particular time instance t :

$$d_{ij,t} = \sqrt{(x_i - x_{j,t})^2 + (y_i - y_{j,t})^2 + (z_i - z_{j,t})^2} \quad (8)$$

$$\cos(\theta_{ij,t}) = \frac{\bar{N} \cdot \bar{L}}{\|\bar{N}\| \|\bar{L}\|} \quad (9)$$

$$I_{R_{si,t}} = \sum_{j=1}^k \frac{I_{lj,t}}{d_{ij,t}^2} \cos(\theta_{ij,t}) \quad (10)$$

$$i = 1, \dots, n$$

By the same derivation, equation (10) can be expanded to equation (11), then rewritten as (12).

$$(I_{M_{si,t}} + B_i + \varepsilon_{i,t}) = \sum_{j=1}^k \frac{I_{lj,t}}{d_{ij,t}^2} \cos(\theta_{ij,t}) \quad (11)$$

$$\varepsilon_{i,t} = \sum_{j=1}^k \frac{I_{lj,t}}{d_{ij,t}^2} \cos(\theta_{ij,t}) - I_{M_{si,t}} - B_i \quad (12)$$

It is easy to identify the necessary conditions to solve this system of nonlinear equations under the assumption that we need at least as many equations as there are variables. The unknowns (or variables) include the coordinates (x_j , y_j , z_j) of all k light sources and their corresponding intensities (I_j) and the systematic biases (B_i) of all n sensors. Therefore, given k light source, n sensors, and m time steps, the total number of unknown variables is $(4k*m+n)$. For each time step, four new variables (x_j , y_j , z_j , I_j) are added to the system. However, the biases for n sensors remain the same for all time steps. At the same time, for each time step, exactly one equation can be written for each of the n sensors, therefore, there are total $(m*n)$ equations in the system. The condition specified in (13) must hold true for this system to have solutions. Rewriting inequality (13) to express n in terms of k and m yields inequality (14). Note that in order to obtain high-quality solution, usually more number of equations than number of variables is needed due to the nonlinear nature of the formulation instance and the presence of random noise.

$$(4k \cdot m + n) \leq (m \cdot n) \quad (13)$$

$$n \geq \frac{4km}{m-1} \quad (14)$$

5.2 Experimental Results

Figure 6 through Figure 8 show the normalized accuracy when there are a total of 1, 2, and 3 light sources present in the environment. We randomly selected n points in the environment, then applied the light source solution solved by non-calibrated and the calibrated sensors to estimate what the light intensity was at each of these n points. We then compared both solutions to the true intensity at these n points. In our experiments, we chose $n = 100$. One horizontal axis indicates different number of sensors in the environment; the other horizontal axis shows various number of time steps considered when the instance is solved. The normalized accuracy is obtained by normalizing the intensity accuracy predicted for the n points by calibrated sensors against the solutions by non-calibrated sensors with respect to the true intensities. More specifically,

$$\frac{\sum_{i=1}^n |I_{i,C} - I_R|}{\sum_{i=1}^n |I_{i,NC} - I_R|} \quad i = 1, \dots, n \quad (15)$$

$I_{i,C}$ is the intensity predicted for point i by the calibrated sensors; $I_{i,NC}$ is the intensity predicted for point i by the non-calibrated sensors; and I_R is the real intensity at point i . Note that as this ratio approaches to 0, it has meaning the calibration process is effective compensating the bias; if the ratio goes above 1, it has meaning the calibration process actually makes the bias worse.

All simulations are conducted under the condition of 25% systematic bias and random noise that follows the Gaussian distribution with mean 0 and sigma 3.

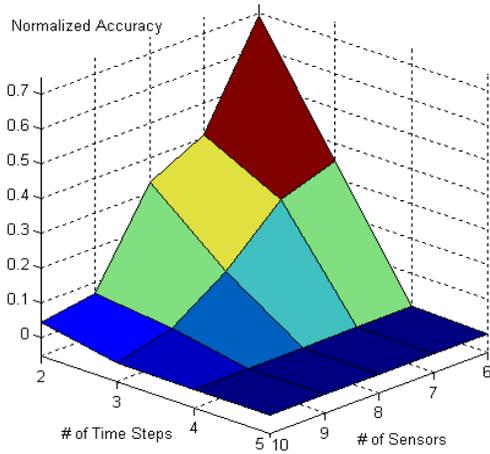


Figure 6. Normalized calibration accuracy for the case of 1 light source

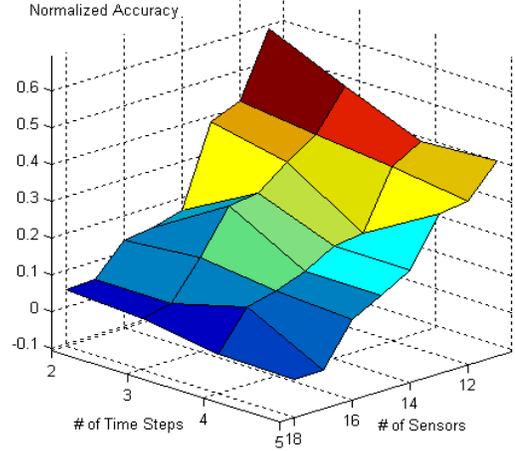


Figure 7. Normalized calibration accuracy for the case of 2 light sources

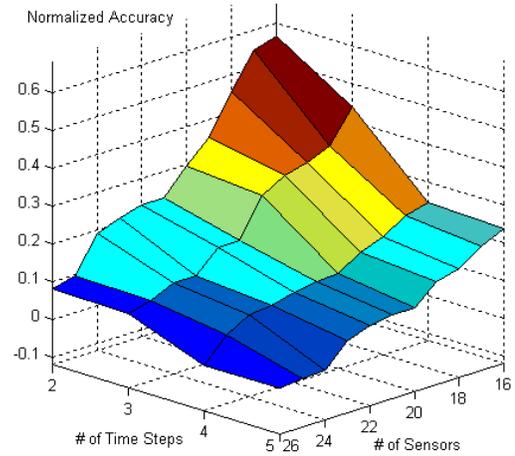


Figure 8. Normalized calibration accuracy for the case of 3 light sources

5.3 Solution Evaluation

The confidence interval can be obtained in two ways. In simulation environment when the true correct location and intensity are available, the confidence interval is obtained by constructing histograms with the center equal to the true location and intensity. However, even in the real-world experiments, confidence interval still can be conducted. It is obtained by applying the resubstitution method where we solve a number of different instances created by considering only a randomly selected subset of readings. As an example, Figure 9 and Figure 10 show the interval of confidence histogram for light source location and intensity error respectively, when 1 light, 8 sensors are present in the environment with 2 time steps. As Figure 9 indicates, 60% of the time, light source location determined by the calibrated sensors are within the interval of $[\text{correct_location} \pm 2.5]$. In Figure 10, with roughly 60% of confidence, the light source intensity calculated by the calibrated sensors is within the interval of $[\text{correct_intensity} \pm 1.7]$.

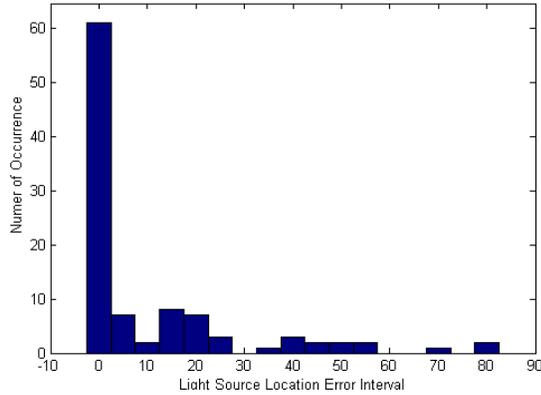


Figure 9. Interval of confidence histogram of light source location

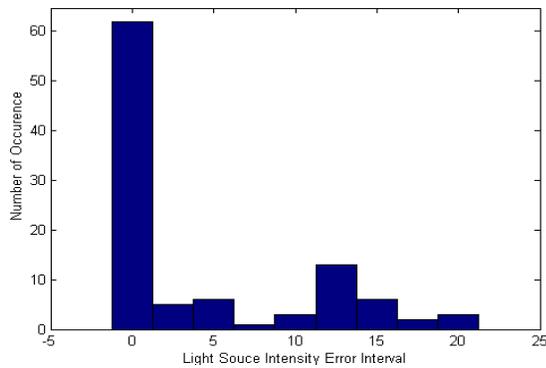


Figure 10. Interval of confidence histogram of light source intensity

6. SUMMARY

We introduced two approaches for calibrating wireless sensor network data. Off-line calibration takes the advantage of the use of high-quality high-cost meters as standards to calibrate inaccurate sensors. We used piece-wise least-squares method to obtain the best fitting mapping function. On-line calibration can be organized within sensors themselves without the existence of standards. More specifically, it does not require additional accurate instruments (or sensors) present. On-line calibration is formulated as a nonlinear function minimization instance and solved by applying the conjugate gradient-based solver. In addition, we applied non-parametric statistical methods to obtain confidence intervals for all results. Both techniques are demonstrated on light sensors.

7. ACKNOWLEDGEMENT

This material is based upon work supported in part by the National Science Foundation under Grant No. ANI-0085773 and NSF CENS Grant.

8. REFERENCES

[1] Bazaraa M., Shetty C.M., and Sherali H., *Nonlinear Programming: Theory & Applications*. Wiley, 1994.

[2] Belsley D.A., Kuh E., and Welsch R.E. *Regression Diagnostics*. New York, Wiley, New York, 1980.

[3] Bertsekas D.P. *Nonlinear Programming*, Second Edition. Athena Scientific, 1999.

[4] Borenstein J. and Feng L. "Measurement and Correction of Systematic Odometry Errors in Mobile Robots". *IEEE Transactions on Robotics and Automation*, 12(6):869-880, 1996.

[5] Chatterjee S. and Price B. *Regression Analysis by Example*. Wiley, New York, 1977.

[6] Davison A.C. and Hinkley D.V. *Bootstrap Methods and their Application*. Cambridge University Press, Cambridge, UK, 1997.

[7] Elson J. and Estrin D. "Time Synchronization for Wireless Sensor Networks". *Tech. Rep. UCLA-CS-020008*, University of California, Los Angeles, May 2002.

[8] Efron B. "Bootstrap Methods: Another Look at the Jackknife". *Annals of Statistics*, 7:1-26, 1979.

[9] Efron B. and Tibshirani R. *An Introduction to the Bootstrap*, Chapman-Hall, 1993.

[10] Faugeras O.D., Luong Q.-T., and Maybank S.J. "Camera Self-calibration: Theory and Experiments". *The International Journal of Computer Vision*, 4(3):225-246, 1992.

[11] Good P.I. *Resampling Methods*, Springer-Verlag, New York, NY, 1999.

[12] Hightower J., Vakili C., and Borriello G. "Design and Calibration of the SpotON Ad-Hoc Location Sensing System". *Unpublished*, August 2001.

[13] Hoff B. and Azuma R. "Autocalibration of an Electronic Compass in an Outdoor Augmented Reality System". *Proceedings of IEEE and ACM International Symposium on Augmented Reality*. Munich, Germany, Oct 2000, pages 159-164.

[14] <http://www-unix.mcs.anl.gov/otc/Guide/faq/nonlinear-programming-faq.html>

[15] <http://www.willnaylor.com/wplib.html>

[16] Kahaner D., Moler C., and Nash S. *Numerical Methods and Software*. Prentice Hall, Englewood Cliffs, New Jersey, 1989.

[17] Rappaport T. *Wireless Communications*, Prentice Hall, New Jersey, 1996.

[18] Press W., Flannery B., Teukolsky S., and Vetterling W., *Numerical Recipes*. Cambridge, 1986.

[19] Quenouille M.H. "Notes on Bias in Estimation". *Biometrika*, 43:353-360, 1956.

[20] Bychkovskiy V., Megerian S., Estrin D., and Potkonjak M. "Colibration: A Collaborative Approach to In-Place Sensor Calibration". *2nd International Workshop on Information Processing in Sensor Networks (IPSN'03)*, Palo Alto, April 2003, pages 301-316.

[21] Welch G. and Bishop G. "SCAAT: Incremental Tracking with Incomplete Information". *Proceedings of ACM SIGGRAPH*, Los Angeles, Aug 1997, pages 333-344.

[22] Whitehouse K. and Culler D. "Calibration as Parameter Estimation in Sensor Networks". *Proceedings of ACM International Workshop on Wireless Sensor Networks and Applications (WSNA'02)*, Atlanta, 2002, pages 59-67.

[23] Wong J., Megerian S., and Potkonjak M. "Design Techniques for Sensor Appliances: Foundations and Light Compass Case Study". *ACM/IEEE Design Automation Conference*, Los Angeles, June 2003, pages 66-71.