Sensor Planning for Range Cameras via a Coverage Strength Model

Jose Luis Alarcon Herrera, Aaron Mavrinac, and Xiang Chen

Abstract— **A method for sensor planning based on a previously developed coverage strength model is presented. The approach taken is known as generate-and-test: a feasible solution is predefined and then tested using the coverage model. The relationship between the resolution of the imaging system and its performance is the key component to perform sensor planning of range cameras. Experimental results are presented; the inverse correlation between coverage performance and measurement error demonstrates the usefulness of the model in the sensor planning context.**

I. Introduction

Sensor planning toward optimal camera placement is an important aspect of system integration in machine vision. The goal of sensor planning is to improve the performance of the vision system, the performance herein is defined as the ability of the system to repeatedly complete a task under controlled conditions. Several methods have been proposed to solve this problem. Typically a set of feasible camera configurations is defined as well as some metric for performance; optimal camera placement is generally achieved by maximizing coverage. This is particulary useful for large multi-camera configurations. However, some machine vision tasks such as industrial inspections require monocular systems and rely more on an approach that takes in to account the task's parameters in detail. As discussed in Mavrinac et al. [1] a global view of the system where coverage is defined as a bivalent condition of visibility is not sufficient; points in the field of view can be fully covered, partially covered or not covered at all, therefore, to express the vagueness of coverage the model assigns coverage strength a value in the range $[0, 1]$. Our task of three-dimensional measurement based on laser scanners is used primarily in industrial inspections where the parameters involved in the scene are strictly controlled, (e.g. no external occlusion is allowed, etc.). Our previously developed coverage model [2], [1] is well suited to a generate-and-test approach. Our coverage metric has been shown to closely reflect the task's a posteriori performance [2], [3], [1]. Currently there exists no feasible technique for numerical optimization using this model; in this paper we employ the generate-and-test approach to perform sensor planning. The main purpose of the current brief is to test the usefulness of the coverage strength model in the sensor planning context. The experimental results are expected to provide preliminary effort toward optimal sensor placement; this is mentioned later in Section VI.

This research was supported in part by the National Council of Science and Technology of Mexico (CONACyT) and by the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors would like to acknowledge the support provided by Vista Solutions Inc.

J. L. Alarcon Herrera, A. Mavrinac and X. Chen are with the Department of Electrical & Computer Engineering, University of Windsor, 401 Sunset Ave., Windsor, Ontario, Canada, N9B 3P4. {alarconj,mavrin1,xchen}@uwindsor.ca

In sensor planning and optimal camera placement, static occlusion and dynamic occlusion present an issue for the maximization of coverage; objects in the scene occlude points of interest, thus preventing the cameras from imaging the entire scene. Dynamic occlusion has been handled using a probabilistic model by Mittal and Davis [4] and Chen and Davis [5]. In the context of laser-based systems, the work of Pito [6] also deals with occlusion, approaching the next-best-view problem by focusing on minimizing occlusion.

As shown in the work of Scott et al. [7] maximizing coverage also involves achieving certain degree of overlap for the case of n-ocular tasks such as surface modeling and reconstruction. In more recent work Scott [8] models the laser scanner system in detail. Prieto et al. [9] give special attention to the effects of the angle between the laser plane and the optical axis of the camera. However, the authors do not include the effects of focal length and aperture diameter in the estimation of good camera placement.

Sensor planning requires a priori information of the system such as camera parameters that allow the computation of some performance metric. A performance metric is then used to assign some meaningful value to a particular camera configuration before it can be selected as a good configuration. Ram et al. [10] developed a performance metric considering such factors as direction of view and zoom. However, the authors neglect distortion caused by perspective projection. Erdem and Sclaroff [11] propose the use of a more realistic model for coverage. The work of González-Banos et al. [12] is more concerned with the accurate representation of performance. In a laser based task, the authors parameterize visibility using conditions such as direction of view and range within the working distance of the camera. Other examples are found in the work of Angella et al. [13] and Hörster et al [14].

The sensor planning literature shows different ways in which coverage is modeled and parameterized; however, most existing models are bivalent and do not always encapsulate all the parameters related to the overall description of coverage. Some models are concerned only with direction of view and zoom such as that of Ram et al. [10], Reed and Allen [15] provide an excellent example, working to solve the next-best-view problem, they consider not only visibility but resolution and direction as well. Their work is also an example of the generate-and-test approach.

This paper is organized as follows. in Section II, we give an overview of the camera parameters and some concepts that are relevant to our task. In Section III, we build the necessary background and describe the coverage strength model: we review the components of the model that account for the various factors involved in the camera's performance. Section V describes the experimental setup and presents the results. Finally, we present some concluding remarks and notes on future work in Section VI.

II. Camera Parameters

The model of a camera has two types of parameters: intrinsic and extrinsic. The intrinsic parameters include the focal length, the effective aperture diameter, the radial distortion coefficients, the physical pixel size, the sensor size in pixels, and the pixel coordinates of the optical center. The extrinsic parameters express the camera position and orientation relative to a reference frame.

Most sensor planning research has proposed methods and algorithms for finding good camera configurations by choosing a solution space over the extrinsic parameters of the camera (which can be continuous or discrete) and optimizing the configuration. In this paper we aim to modify not only the extrinsic parameters but also the intrinsic parameters through the use of a realistic coverage strength model (see Section III), that takes into account all of the aforementioned characteristics of the camera to achieve an accurate description of coverage.

A. Camera Calibration

A laser-based 3D imaging system is typically configured as shown in Figure 1. The two main characteristics of this configuration are the camera and the laser plane. Both the coverage strength model and the 3D measurement algorithm rely on the camera's parameters, thus camera calibration is necessary.

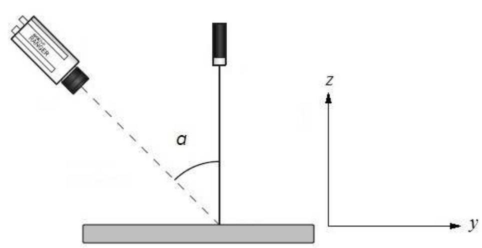

Fig. 1. Typical Camera Setup

The calibration procedure comprises two stages. The first corrects for lens distortion where image coordinates $(u^{'}, v^{'})$ are calculated from the raw pixel coordinates (u, v) using Brown's lens distortion model [16].

$$u^{'} = u + u_o(C_1 r^2 + C_2 r^4) + 2C_3 u_o v_o + C_4 r^2 + 2u_o^2 \quad (1)$$

$$v^{'} = v + v_o(C_1 r^2 + C_2 r^4) + 2C_4 u_o v_o + C_3 r^2 + 2v_o^2 \quad (2)$$

$$u_o = u - o_u \quad v_o = v - o_v \quad r = \sqrt{u_o^2 + v_o^2} \quad (3)$$

where (o_u, o_v) are the pixel coordinates of the image projection of the optical center and C_1 to C_4 are the lens distortion coefficients.

The second stage produces a homography between the two-dimensional image plane and the two-dimensional laser plane defined in homogeneous coordinates as a 3×3 matrix $\mathbf{H}$:

$$\begin{bmatrix} x \\ z \\ s \end{bmatrix} = \mathbf{H} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (4)$$

where s is a scale factor.

Mavrinac et al. [17] provide the derivation and implementation details of this calibration procedure.

B. Measurement Resolution and Occlusion

In this paper we consider two types of resolution: first, the optical resolution which is the ability of the camera to capture in detail the object in the field of view; this is defined by the sensor's pixel size in micrometers together with the number of pixels needed to form a feature in the image. Second, the measurement resolution, (also known as height resolution), refers to the minimum change in the position of the laser line that can be detected by the camera along the z axis.

As discussed in Section V-C, in order to choose good camera placement we extend the coverage model to account for measurement resolution. As will be made clear, the addition of this factor yields a more accurate description of coverage which is closely related to the a posteriori performance of the task.

Sensor planning in laser based tasks is directly related to the accuracy of the measurements and the completeness of the image; performance is the degree of accuracy and coverage that the system is is able to achieve. The performance of laser scanner tasks is negatively affected by two types of occlusion: laser occlusion and camera occlusion [18]. The first occurs when the laser is unable to illuminate a point in the object that needs to be visible from the camera, this is generally the case for non-convex shapes. The second takes place when the camera is unable to image the scene due to self-occlusion of the object of interest. Occlusion is not addressed in this paper and is left as subject for future work.

III. Coverage Strength Model

In previous work, Mavrinac et al. [2], [3], [1] developed a coverage strength model which includes most of the camera's characteristics and properties; among these are the extrinsic and intrinsic parameters as well as the optical properties of the lens, the camera's sensor and several intuitive task parameters which will be described in this section.

A. General Model

The coverage strength model of a given camera system assigns to every point in the stimulus space a measure of coverage.

Definition 1: The three-dimensional directional space $\mathbb{D}^3 = \mathbb{R}^3 \times [0, \pi] \times [0, 2\pi)$ consists of three-dimensional Euclidean space plus direction, with elements of the form (x, y, z, ρ, η).

Definition 2: A coverage strength model is a mapping $C : \mathbb{D}^3 \rightarrow [0, 1]$, for which $C(\mathbf{p})$, for any $\mathbf{p} \in \mathbb{D}^3$, is the strength of coverage at $\mathbf{p}$.

Definition 3: A relevance model is a mapping $R : \mathbb{D}^3 \rightarrow [0, 1]$, for which $R(\mathbf{p})$, for any $\mathbf{p} \in \mathbb{D}^3$, is the minimum desired coverage strength or coverage priority at $\mathbf{p}$.

We term $\mathbf{p} \in \mathbb{D}^3$ a *directional point*. For convenience, we denote its spatial component $\mathbf{p}_s = (\mathbf{p}_x, \mathbf{p}_y)$ or $\mathbf{p}_s = (\mathbf{p}_x, \mathbf{p}_y, \mathbf{p}_z)$ and its directional component $\mathbf{p}_d = (\mathbf{p}_\rho, \mathbf{p}_\eta)$. $\eta \in [0, 2\pi)$, $\rho \in [0, \pi]$.

The coverage performance of a sensor system is given by

$$m(C, R) \equiv \frac{|\dot{C} \cap \dot{R}|}{|\dot{R}|} \tag{5}$$

where $\dot{C}$ is the coverage strength model sampled on a discrete grid of points in $\mathbb{D}^3$, similarly, $\dot{R}$ is the discretized relevance model.

Here we detail the coverage strength model parametrization for cameras, which consists of four components: visibility, resolution, focus, and direction (angle of view). We omit most of the derivation as this is covered in previous work [1]. Throughout, $B_{[0,1]}(x) = \min(\max(x, 0), 1)$ is a function that limits the value x to $[0, 1]$.

The first component, C_V, characterizes visibility. The pinhole camera model is used to compute the angles of the field of view of the camera. A task parameter γ is introduced to account for the partial coverage of non-point features located near the boundaries of the field of view. γ is measured in pixels and it reflects the expected size of the feature's neighborhood. The horizontal and vertical cross-sections are given by

$$C_{Vh}(\mathbf{p}) = B_{[0,1]} \left(\frac{\min\left(\frac{\mathbf{p}_x}{\mathbf{p}_z} + \sin(\alpha_{hl}), \sin(\alpha_{hr}) - \frac{\mathbf{p}_x}{\mathbf{p}_z}\right)}{\gamma_h} \right) \tag{6}$$

$$C_{Vv}(\mathbf{p}) = B_{[0,1]} \left(\frac{\min\left(\frac{\mathbf{p}_y}{\mathbf{p}_z} + \sin(\alpha_{vt}), \sin(\alpha_{vb}) - \frac{\mathbf{p}_y}{\mathbf{p}_z}\right)}{\gamma_v} \right) \tag{7}$$

where α_{hl}, α_{hr} are the horizontal angles of view, and α_{vt} and α_{vb} are the vertical angles of view, of a rectilinear projected image, γ_h and γ_v are the horizontal and vertical offsets calculated from γ (see Mavrinac et al.[1]).
the complete C_V is given by

$$C_V(\mathbf{p}) = \begin{cases} \min(C_{Vh}(\mathbf{p}), C_{Vv}(\mathbf{p})) & \text{if } \mathbf{p}_z > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

The second component, C_R, characterizes pixel resolution (number of pixels per unit distance). The resolution is a function of the distance between a point and the principal point along the optical axis. Two task parameters are introduced; R_1 which is the ideal pixel resolution and R_2 which is the minimum resolution. The resolution component is given by

$$C_R(\mathbf{p}) = B_{[0,1]} \left(\frac{z_2 - \mathbf{p}_z}{z_2 - z_1} \right) \tag{9}$$

for $R_1 > R_2$, where the values of z_1 and z_2 are given by (10), substituting task parameters R_1 and R_2, respectively, for R.

$$z_R = \frac{1}{R} \min \left[\frac{w}{2\sin(\alpha_h/2)}, \frac{h}{2\sin(\alpha_v/2)} \right] \tag{10}$$

The third component, C_F, characterizes focus (depth of field). The task parameter $c_{\max}$ indicates the maximum blur circle diameter that can be tolerated.

The component C_F is given by

$$C_F(\mathbf{p}) = B_{[0,1]} \left(\min \left(\frac{\mathbf{p}_z - z_n}{z_\triangleleft - z_n}, \frac{z_f - \mathbf{p}_z}{z_f - z_\triangleright} \right) \right) \tag{11}$$

where the values of z_n and z_f are the near and far limits of the depth of field as given by (12) substituting $c_{\max}$ for c. Similarly substituting $c_{\min}$ for c in (12) yields $z_\triangleleft$ and $z_\triangleright$.

$$z = \frac{Afz_S}{Af \pm c(z_S - f)} \tag{12}$$

The fourth component, C_D, characterizes direction (angle of view). A point p is visible if the camera lies in the half-space defined by the plane tangent to the surface of the point; p is visible from the camera only if

$$\Theta(\mathbf{p}) \equiv \mathbf{p}_\rho - \left(\frac{\mathbf{p}_y}{r} \sin \mathbf{p}_\eta + \frac{\mathbf{p}_x}{r} \cos \mathbf{p}_\eta \right) \arctan \left(\frac{r}{\mathbf{p}_z} \right) \geq \frac{\pi}{2} \tag{13}$$

where $r = \sqrt{\mathbf{p}_x^2 + \mathbf{p}_y^2}$.

The task parameters ζ_1 and ζ_2 are the ideal and maximum angles between the normal of a feature and the optical axis. The direction component is given by

$$C_D(\mathbf{p}) = B_{[0,1]} \left(\frac{\Theta(\mathbf{p}) - \pi + \zeta_2}{\zeta_2 - \zeta_1} \right) \tag{14}$$

The full model is given by

$$C(\mathbf{p}) = C_V(\mathbf{p}) C_R(\mathbf{p}) C_F(\mathbf{p}) C_D(\mathbf{p}) \tag{15}$$

IV. Sensor Planning

The following iterative procedure is an application of the coverage strength model, the objective is to select good camera configurations from a feasible solution space.

1. Based on the geometry of the scene, predefine a discrete solution space: from an initial configuration, iteratively change the camera parameters. For simplicity these are categorized as position, orientation, and intrinsic parameters.
2. Define the relevance model for the task. In this case the relevance model is a discetized subset of the laser plane within the operational field of view of the camera, and it is in the same reference frame.
3. Select a camera configuration from the solution space.
4. Compute the coverage strength of the selected configuration.
5. Repeat steps 3 and 4 over the solution space and output the configuration with the highest coverage performance.

Moreover, if it is desired to change the camera configuration, such as a change in the optics (i.e. aperture, focal length); this model facilitates the investigation of the effect on the performance of the imaging system; thus, the need to make any physical changes is eliminated.

V. Experimental Results

A. Apparatus

In the experiments, the camera used is the SICK-IVP Ranger E industrial 3D camera with a laser line projector. The camera and laser were mounted and calibrated using two different calibration techniques: laser line calibration [17] and full camera calibration [19].

Laser line calibration is used to produce the lookup table required by the ranger to output calibrated images. Calibrated images have pixel values given in millimeters with respect to the reference frame defined during calibration.

Full calibration[1] is used to compute the camera's intrinsic and extrinsic parameters that are required to generate the coverage model. The laser line calibration generates a mapping from a two-dimensional plane to a two-dimensional plane, from this mapping the three-dimensional pose of the camera cannot be estimated directly. Moreover the current laser line calibration method does not compute the focal length, therefore full calibration is also required in addition to the look-up-table generation. Lens distortion was corrected in all the experiments.

With the system mounted and calibrated, several pictures of the target were taken using different camera-laser configurations. The system had to be calibrated every time the camera and laser where rearranged. Different target positions were used in order to cover most of the field of view of the camera and thus not limit the results to a particular case; the accuracy of the measurements is not the same everywhere in the field of view; this is caused by perspective projection.

The target is the calibration object used for laser calibration, it has a series of triangles of known dimensions. The image processing software developed in HALCON takes as an input the calibrated image generated using the look-up-table available from calibration, then, the software detects the triangles on the image and measures the height.

B. Software

In this experiment, we use our Adolpus[2] simulation software to compute the coverage performance for a particular camera configuration. The model is parameterized using the camera system. The intrinsic and extrinsic parameters used in the model are those of the physical system. Most of the image processing and calibration is performed using the HALCON machine vision libraries [20].

The camera system was estimated as shown in the simulation example in Figure 2

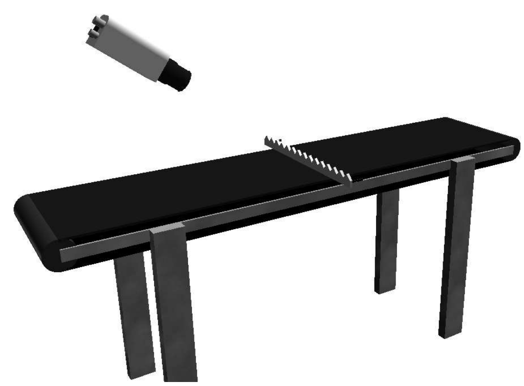

Fig. 2. Software Simulation

C. Task-Related Parametrization

Range cameras are very robust to blur due to de-focus; when a profile is acquired by the camera and the laser line is extracted, the camera computes the center of gravity, allowing for high accuracy even with an image that is out of focus. The focus parameter was set to a relatively large value; blur circles of 1.0 mm are the maximum blur allowed.

The relevance model in this experiment is defined as the points of interest of the target; which are the crests and valleys of the triangular shape in the calibration target. The points of interest are directional points in $\mathbb{R}^3$, with direction normal to the face of the features themselves; in other words the direction is parallel to the y axis. (see Figure 1).

The ideal angle for best resolution is $\zeta_1 = 0$; the resolution of the camera increases as the angle α, (see Figure 1), increases until the optical axis of the camera becomes orthogonal to the laser plane [18]. The second parameter was selected as the angle at which the camera can no longer collect any useful information, so $\zeta_2 = \frac{\pi}{2}$.

[1]Performing two different types of calibration in this paper is only necessary for the experiments and is not normally required for sensor planning

[2]Adolphus is free software licensed under the GNU General Public License. Complete Python source code and documentation are available at `http://github.com/ezod/adolphus`

The parameters γ, R_1 and R_2 are measured in pixels. An estimated size of the features detected by the image processing software was selected as the value for γ, so $\gamma = 6$, similarly $R_1 = 5.22$ and $R_2 = 1.0$ were selected as the ideal and minimum cutoff resolutions, respectively.

D. Results and Analysis

Using the measurement software, the data was compared with the ground truth; this is to establish the performance of the physical system. The measured performance is then compared with the coverage strength. As an example, four examples from the experiment data pool are shown in table I.

TABLE I
COVERAGE STRENGTH AND THE MEASUREMENT ERROR

Camera	α	Coverage Strength	Measurement Error (mm)
C1	53.88°	0.6122	0.1472
C2	51.18°	0.5706	0.4498
C3	34.46°	0.3275	0.8167
C4	16.91°	0.1859	0.9311

where α is the angle between the laser plane and the optical axis of the camera, as shown in Figure 1.

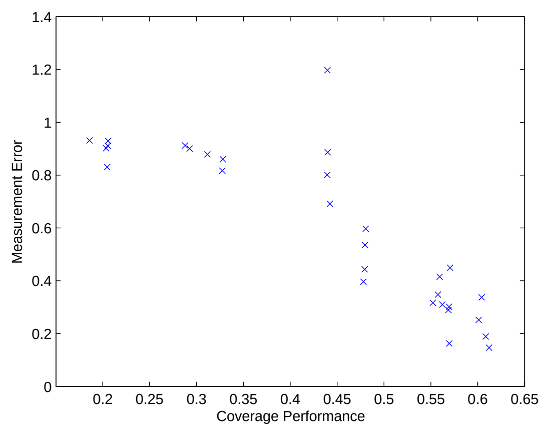

Fig. 3. Performance Correlation

The Pearson correlation coefficient is calculated between the coverage metric and the measurement error. The correlation is $r = -0.8508$.

The predicted performance of the system measured throughout the coverage strength is closely related to the performance of the task, it is clear that choosing camera configuration number one from the example in table I will yield the most accurate results. Sensor planning is then possible by predefining a set of feasible camera configurations and then computing the coverage strength to select the one with the highest value.

VI. CONCLUSIONS AND FUTURE WORK

A. Conclusions

The sensor planning task for the case of visual sensors can be achieved through the selection of best-camera configuration based on the information provided by the coverage strength model. Moreover the coverage strength model can be easily adapted according to the needs of the task. It has been shown that the model is flexible to another kind of task: three-dimensional measurement where the model was adapted to account for the height resolution.

B. Future Work

As described in section II-B, laser scanners are highly affected by camera occlusion; the laser light is being blocked by the object of interest which is not known a priori. This can be seen as dynamic occlusion which is hard to predict and include in the coverage metric. One way to approach this in future work is to develop a probabilistic model for camera occlusion. Another, and more interesting subject is to find a suitable method for optimal sensor placement; exact solutions are not feasible because of the computational cost as explained by Hörster et al. [14], the challenge is then to find the best approximation.

REFERENCES

[1] A. Mavrinac, X. Chen, and J. L. Alarcon-Herrera, "Coverage Strength Models for Multi-Camera Systems," *IEEE Trans. Pattern Analysis and Machine Intelligence*, submitted for publication.

[2] A. Mavrinac, J. L. Alarcon-Herrera, and X. Chen, "A Fuzzy Model for Coverage Evaluation of Cameras and Multi-Camera Networks," in *Proc. 4th ACM/IEEE Intl. Conf. on Distributed Smart Cameras*, 2010.

[3] ——, "Evaluating the Fuzzy Coverage Model for 3D Multi-Camera Network Applications," in *Proc. Intl. Conf. on Intelligent Robotics and Applications*, 2010.

[4] A. Mittal and L. Davis, "A General Method for Sensor Planning in Multi-Sensor Systems: Extension to Random Occlusion," *International Journal of Computer Vision*, vol. 76, pp. 31–52, 2008.

[5] X. Chen and J. Davis, "An Occlusion Metric for Selecting Robust Camera Configurations," *Machine Vision and Applications*, vol. 19, pp. 217–222, 2008.

[6] R. Pito, "A Solution to the Next Best View Problem for Automated Surface Acquisition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, no. 10, pp. 1016–1030, 1999.

[7] W. Scott, G. Roth, and J.-F. Rivest, "View Planning With a Registration Constraint," in *3-D Digital Imaging and Modeling. Proceedings. Third International Conference on*, 2001, pp. 127–134.

[8] W. Scott, "Model-Based View Planning," *Machine Vision and Applications*, vol. 20, pp. 47–69, 2009.

[9] F. Prieto, R. Lepage, P. Boulanger, and T. Redarce, "A CAD-Based 3D Data Acquisition Strategy for Inspection," *Machine Vision and Applications*, vol. 15, pp. 76–91, 2003.

[10] S. Ram, K. R. Ramakrishnan, P. K. Atrey, V. K. Singh, and M. S. Kankanhalli, "A Design Methodology for Selection and Placement of Sensors in Multimedia Surveillance Systems," in *Proceedings of the 4th ACM International Workshop on Video Surveillance and Sensor Networks*, 2006, pp. 121–130.

[11] U. M. Erdem and S. Sclaroff, "Automated Camera Layout to Satisfy Task-Specific and Floor Plan-Specific Coverage Requirements," *Computer Vision and Image Understanding*, vol. 103, no. 3, pp. 156–169, 2006.

[12] H. González-Banos and J. C. Latombe, "A Randomized Art-Gallery Algorithm for Sensor Placement," in *Proc. 17th Ann. Symp. on Computational Geometry*, 2001, pp. 232–240.

[13] F. Angella, L. Reithler, and F. Gallesio, "Optimal Deployment of Cameras for Video Surveillance Systems," in *Proc. IEEE Conf. on Advanced Video and Signal Based Surveillance*, 2007, pp. 388–392.

[14] E. Hörster and R. Lienhart, "On the Optimal Placement of Multiple Visual Sensors," in *Proc. 4th ACM Intl. Wkshp. on Video Surveillance and Sensor Networks*, 2006, pp. 111–120.

[15] M. Reed and P. Allen, "Constraint-Based Sensor Planning for Scene Modeling," in *Computational Intelligence in Robotics and Automation. IEEE International Symposium on*, 1999, pp. 131–136.

[16] D. C. Brown, "Decentering Distortion of Lenses," *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444–462, 1966.

[17] A. Mavrinac, X. Chen, P. Denzinger, and M. Sirizzotti, "Calibration of Dual Laser-Based Range Cameras for Reduced Occlusion in 3D Imaging," in *Proc. IEEE/ASME Intl. Conf. on Advanced Intelligent Mechatronics*, 2010.

[18] *Ranger E/D Reference Manual*, SICK IVP, 206.

[19] Z. Zhang, "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations," in *Computer Vision. The Proceedings of the Seventh IEEE International Conference on*, vol. 1, 1999, pp. 666–673.

[20] HALCON. MVTec Software GmbH. [Online]. Available: http://www.mvtec.com/halcon