Multi-objective Sensor Planning for Efficient and Accurate Object Reconstruction

Enrique Dunn and Gustavo Olague

Centro de Investigación Científica y Educación Superior de Ensenada
División de Física Aplicada, EvoVisión Lab
{edunn,olague}@cicese.mx

Abstract. A novel approach for sensor planning, which incorporates multi-objective optimization principals into the autonomous design of sensing strategies, is presented. The study addresses planning the behavior of an automated 3D inspection system, consisting of a manipulator robot in an Eye-on-Hand configuration. Task planning in this context is stated as a constrained multi-objective optimization problem, where reconstruction accuracy and robot motion efficiency are the criteria to optimize. An approach based on evolutionary computation is developed and experimental results shown. The obtained convex Pareto front of solutions confirms the conflict among objectives in our planning.

1 Introduction

Automated visual inspection is a complex task that has been addressed in different scientific communities such as computer vision, photogrammetry and robotics, see [1], [2], [3]. In fact, sensor planning is an interdisciplinary research field that can benefit from incorporating knowledge from different areas in a complementary manner. One of the challenges, in automating the decision making process for inspection systems, is the development of a sensor planner which incorporates different qualitative aspects (e.g. efficiency, accuracy, robustness) involved in a particular vision task. The literature indicates that research on this field generally studies a particular goal of the overall perception task such as minimizing sensing actions, obtaining complete coverage of interest regions or improving image quality. Consequently, despite the variety of methodologies previously proposed, the multi-objective (MO) nature of such planning has not yet been addressed. The choice of an appropriate framework is crucial to incorporate such study. For instance, approaches based on *Generate and Test* iterations, [4], are dependent on user expertise and interaction. *Complete Enumeration* is by definition inadequate for complex tasks due to its computational burden. The *Synthesis* approach, [5], based on the use of analytical functions, presents difficulties to incorporate the MO framework due to the complexity of obtaining an analytical multi-objective solution. *Expert systems*, [3], are a promising approach for MO, but concerns about the dependence on expert prior knowledge may hinder their solution quality in comparison to less heuristic based approaches. A promising alternative is the use of *Simulation Techniques*, [6], which using proper

G.R. Raidl et al. (Eds.): EvoWorkshops 2004, LNCS 3005, pp. 312–321, 2004.

mathematical models, allow to evaluate different application scenarios and gain insight into the studied problem. It is within this context that we present a novel approach that introduces the concept of MO optimization into the sensor planning problem. To our knowledge this is the first time that such a study has been undertaken.

This work deals with the automation of a close range photogrammetric system. We consider the use of an anthropomorphic manipulator arm equipped with a CCD camera on its end effector. The decision making process consists on designing a group of sensing actions that lead to the obtainment of a highly accurate three dimensional reconstruction. The problem is stated within the MO optimization framework. Building on previous work on sensor planning [7],[8], a solution is developed under the evolutionary computation (EC) paradigm.

The goal of this paper is two-fold. First, to describe the MO nature of the sensor planning problem for automated visual measurement tasks. Second, to present an approach based on evolutionary computation to solve such planning.

The outline of the paper is the following. Initially, a description of the sensor planning problem is presented. Then, an explanation on the MO nature of sensor planning and the aspects considered in this work is brought forth. Next, the proposed evolutionary computation approach to solve our planning problem is detailed. This is followed by experimental results. Finally, conclusions and future research guidelines are presented to end the paper.

2 Problem Statement

In the context of our automated visual inspection system, a sensor planner must explicitly specify sensing task execution. Sensing actions are carried out by a physical mechanism Q, in accordance to a set of sensing viewpoints V. Each viewpoint is a vector parametrization $V_i = [v_1, \ldots, v_s]$ which describes a single camera position (X, Y, Z), orientation angles (ω, ϕ, κ), as well as any of the configurable camera intrinsic parameters. The configuration of a manipulator consisting of r rotational joints can be represented as vector $Q = [q_1, \ldots, q_r]$. Given prior knowledge of its Denavit-Hartenbeg parametrization D, this information is sufficient to determine the position and orientation of the end effector where the camera is mounted. Moreover, a viewpoint specification can be obtained by a function of the form $\Gamma(Q, D) \in \mathbb{R}^s$.

A sensing plan consists in a motion trajectory, $Q(t)$, to be followed by the manipulator which executes sensing actions at n different viewpoints V_i. In this work a viewpoint based task specification is utilized.

Definition 1. Viewpoint Based Task Specification: *Obtain a set V of viewpoints $V_i \in \mathbb{R}^s$ from which to determine a time parametrized robot motion trajectory $Q(t)$ where $\forall i \in [1, 2, \ldots, n] \exists \Gamma(Q(t_i), D) = V_i$*

Given some environmental description E, and depending on the chosen task T, a set of operational restrictions $g_i(\cdot) \geq 0$ for $i = 1, \ldots, m$ can be evaluated. Assuming there is parametrization of the set of V into a set of l parameters, the sensor planning problem can be stated as

Definition 2. Multi-objective Sensor Planning: *Find the solution vector* $\boldsymbol{x}^* = [x_1^*, x_2^*, \ldots, x_l^*]^t$ *which satisfies the* m *environmental and task defined constraints:* $g_i(\boldsymbol{x}, E, T) \geq 0 \quad i = 1, 2 \ldots, m$*; adheres to the variables bounds:* $x^{(L)} \leq x_i \leq x^{(U)}$ *and optimizes the vector function* $\boldsymbol{f}(\boldsymbol{x}) = [f_1(\boldsymbol{x}), \ldots, f_k(\boldsymbol{x})]^t$.

The nature of the different objective functions $f_i(\boldsymbol{x})$ and of the problem constraints $g_i(\boldsymbol{x}, E, T)$, shall be given in subsequent sections of this paper. Under the previous definition, the specified constraints define the feasible region $\Omega = \boldsymbol{x} \in \mathbb{R}^l$ of the decision variable space. Therefore, any point in Ω defines a feasible solution. The vector function $\boldsymbol{f}(\boldsymbol{x})$ maps the set Ω into a set $\Phi = \boldsymbol{y} \in \mathbb{R}^k$ that represents all possible values of the objective functions. The set of optimal solutions for a MO optimization problem consists of all decision vectors for which the corresponding objective vectors cannot be improved in any dimension without degradation of another. Mathematically, the concept is stated as follows:

Definition 3. Pareto Optimality: *A point* $\boldsymbol{x}^* \in \Omega$ *is Pareto optimal if for every* $\boldsymbol{x} \in \Omega$ *and* $I = 1, 2, \ldots, k$ *either,* $\forall_{i \in I}(f_i(\boldsymbol{x}) = f_i(\boldsymbol{x}^*))$ *or, there is at least one* $i \in I$ *such that* $f_i(\boldsymbol{x}) > f_i(\boldsymbol{x}^*)$.

The set of optimal task specifications vectors that satisfy such definition is known as the Pareto Optimal Set P^*. These vectors are mapped by $\boldsymbol{f}(\cdot)$ to a subset on the boundary of Φ. These vectors in the objective function space are denominated Pareto Front PF^*.

3 The Multi-objective Nature of Sensor Planning

The motivation for stating sensor planning as MO is based on the characterization of a *"good"* sensing plan. Compliance with the task goals is an obvious requirement. However, we identify two general aspects in assessing the worthiness of a high level task specification : *Solution Quality* and *Process Efficiency*. In highly accurate 3D inspection, solution quality entails minimizing a measure of the 3D reconstruction uncertainty. On the other hand, the process efficiency can be defined in terms of the resources adjudicated to a specific task. The description of these different and usually competing aspects, as well as their integration into our sensor planning approach, is presented next.

3.1 Solution Quality: Accurate 3D Measurements

It is known in the photogrammetric and computer vision communities that viewpoint selection is a determining factor in accurate 3D reconstruction. Designing a multiple viewpoint configuration (e.g. photogrammetric network) is a complex geometrical problem. In fact, the photogrammetric network design search space is multi-modal. Therefore, for a given scenario, there exist several very different spatial distribution of viewpoints which yield similar reconstruction accuracy. In order to derive a useful criteria, an approach based on the error propagation phenomena, as presented by Olague [9], will be utilized. Relying on the implicit

function theorem, the covariance matrix ΛP of a 3D reconstruction can be obtained having knowledge of: 1) the 2D image measurement uncertainty Λp and, 2) a 3D reconstruction model from 2D image data of the form $P = f(p)$. Such relationship is stated as follows.

Proposition 1. *Given a random variable $p \in \mathbb{R}^m$, of Gaussian distribution, mean $E[p]$, and covariance Λp, and $P \in \mathbb{R}^n$, the random vector given by $P = f(p)$, where f is a function of class C^1, the mean of P can be approximated to a first-order Taylor expansion by $f(E[p])$ and its covariance by:*

$$\Lambda P = \frac{\partial f(E[p])}{\partial p} \Lambda p \frac{\partial f(E[p])^t}{\partial p} \tag{1}$$

The 2D uncertainty Λp can be determined as a function of the incidence angle between the camera viewing direction and the surface normal of a 3D point. The function model of this phenomena in each of image axes is obtained using experimental measurements of the projectively invariant cross-ratio. The resulting model is given by the function

$$y = \beta \left(e^{\left(\frac{\alpha}{90-x}\right)} + e^{\left(\frac{\alpha}{90+x}\right)} \right) + \gamma, \tag{2}$$

where the best fit parameters are: α=79.74, $\beta = 1.31 \times 10^{-3}$, and $\gamma = 8 \times 10^{-3}$. Once an analytical expression for the reconstruction covariance matrix is obtained, in order to derive a criteria, a metric for comparing covariance matrices needs to be adopted. For this work, the maximum element of the main diagonal of ΛP was selected as a criteria:

$$f_1(\boldsymbol{x}) = \max_{i=1\ldots3} \Lambda P_{ii} \tag{3}$$

3.2 Process Efficiency: Operational Costs

In an automated iterative process, an active vision system should make optimal use of the manipulator infrastructure. This entails considerations regarding kinematic and dynamic characteristics of a task planning specification, such as: *distance* traveled by the manipulator, total *effort* required for the motion, total *time* required for the movement or collision *risk*.

Such operational costs are evaluated from a complete sensing task specification which depends on viewpoint selection V, as well as the environmental configuration E. A problem specific mapping from these descriptions into an explicit and applicable task plan is needed. We state such mapping as $S(V, E)$, which result can be generally stated as a time parametrized function of the robot joint values (e.g. configuration space), $\boldsymbol{Q}(t) \in \mathbb{R}^6$ where $t \in [a, b]$ in the continuous case or $t \in [1, 2, \ldots, n]$ for the discrete one. Moreover, this general mapping $S(\cdot)$ can be adapted to consider optimal path planning, obstacle avoidance or kinematic and dynamic considerations. Since a general cost function $\Psi(\cdot) \in \mathbb{R}$ evaluates robot motion, such a function can be estimated from two robot configurations obtained at different times during task execution, which

yields $\Psi(\boldsymbol{Q}(t_1), \boldsymbol{Q}(t_2))$. Consequently, in order to evaluate the operational cost over the entire course of task execution we have for the continuous case:

$$f_2(\boldsymbol{x}) = \int_b^a \Psi(\boldsymbol{Q}(t), \boldsymbol{Q}(t+dt))dt, \tag{4}$$

while the discrete case yields

$$f_2(\boldsymbol{x}) = \sum_{t=1}^{n-1} \Psi(\boldsymbol{Q}(t), \boldsymbol{Q}(t+1)). \tag{5}$$

4 Our Evolutionary Computation Approach

Evolutionary algorithms operate over multiple solutions in a concurrent manner, allowing the attainment of several different solutions at the end of a single execution run. This has lead to the design of methodologies that direct the whole population toward the Pareto optimal set. Furthermore, the stochastic nature of the EC approach has showed to be robust against the complexity of the search space, such as the one present in sensor planning [9].

4.1 Solution Representation

In this work the camera placement is restricted to an inward looking viewing sphere centered around the measured object. This simplifies the viewpoint specification, since the viewing direction is implicitly expressed as the sphere center. Moreover, by specifying a sufficient viewing distance from the object, it is possible to satisfy the focus, resolution and field of view restrictions common to viewpoint selection. In virtue of the viewing sphere model, an individual viewpoint can be expressed by its polar coordinates. Using a constant radius, the 3D position of a viewpoint can be specified by a 2D real valued vector $[\alpha, \beta]$. Given an n number of desired images, the constant length representation vector used for specifying an imaging geometry configuration is denoted as

$$\boldsymbol{x} \in \mathbb{R}^{2n} \quad where \quad \alpha_i = x_{2i-1}, \beta_i = x_{2i} \quad for \quad i = 1, \ldots, n. \tag{6}$$

4.2 Constraint Handling

It is important to note that not all points on the surface of the viewing sphere are valid, nor all the possible combinations of viewpoints conform a valid network configuration. This is due to the local and global restrictions involved in our problem. A description of our camera placement search space is given in [7].

Local restrictions are imposed to the viewpoint selection process and can be dependent upon the object studied or the working environment. The case of optical self-occlusion for complex objects limits the visibility region of some of the features on the object to a subset of the viewing sphere. In fact, it is possible for a particular viewpoint not to be able to capture any of the interest points on

the object. Additionally, when considering the incorporation of a manipulator robot into the sensing task, it is necessary to also take into account kinematic restrictions. For example, a given viewpoint may not be reachable to the robot.

Global restrictions are imposed to the final task specification. In photogrammetric reconstruction this translates to data sufficiency requirements. In order to be effective, a network configuration should have a sufficient amount of redundancy in its observations. Therefore, it is possible that a certain configuration does not provide the necessary data for the triangulation of a particular object feature. This data inconsistency leads to very poor overall reconstruction due to the least squares method involved in such a process. Another type of global restrictions are the ones imposed by any specified bounds on the overall performance of the sensing task. For example: maximum displacement allowed or minimum accuracy obtained.

To incorporate the above mentioned constraints different approaches are adopted. Viewpoint selection constraints (i.e. local restrictions) are enforced using a deterministic repair mechanism. This procedure consists in:

1. Initializing an independent random number generator with a value obtained from a fixed linear combination of the α, β values corresponding to the invalid viewpoint.
2. Iteratively generate new values for α, β until a valid viewpoint is generated.
3. Locally replace such repaired values. That is, they are only used for objective function evaluation.

This enables the evolutionary algorithm to work only on valid information during the optimization process, and in turn, restrict the search space. Since the random number generator is initialized in a deterministic manner, each time that the same invalid viewpoint is repaired, the adjusted values will remain constant. Meanwhile, the 3D reconstruction data requirements (i.e. global restrictions) are enforced by means of penalty function assessment. This procedure identifies the cases when data consistency requirements are not meet and assigns an arbitrary value to the object function evaluation. This value is chosen in such a way that the violating individuals still have a possibility to contribute to the evolutionary process (i.e. $f_2(\boldsymbol{x}) = 2.0$). Both of these restriction handling politics are implemented at the objective function evaluation phase, allowing a relative independence from the choice of evolutionary engine used for optimization.

4.3 Genetic Operators

The simulated binary crossover (SBX) emulates the working principle of the single point crossover operator on binary strings. From two parent solutions P_1 and P_2, it creates two children C_1 and C_2 as follows:

$$\begin{aligned} C_1 &= 0.5[(1+\beta)P_1 + (1-\beta)P_2] \\ C_2 &= 0.5[(1-\beta)P_1 + (1+\beta)P_2] \end{aligned} \quad \text{with } \beta = \begin{cases} (2u)^{\frac{1}{\eta_x+1}} & \text{if } u < 0.5 \\ \left(\frac{1}{2(1-u)}\right)^{\frac{1}{\eta_x+1}} & \text{otherwise.} \end{cases} \tag{7}$$

The spread factor β is dependent on a random variable $u \in [0, 1]$ and on an user defined nonnegative value η_x that characterizes the distribution of the

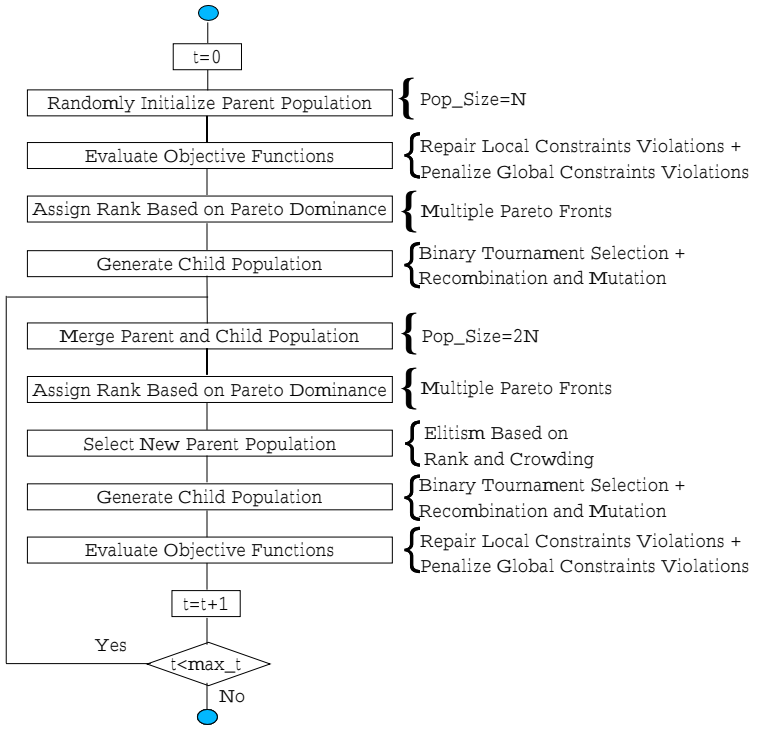

Fig. 1. Flow Diagram of our NSGA-II based approach.

children in relation to their parents. The mutation operation modifies a parent P into a child C using the boundary values $P^{(LOW)}$ and $P^{(UP)}$ of each the decision variables in the following manner:

$$C = P + (P^{(UP)} - P^{(LOW)})\delta \quad \text{with } \delta = \begin{cases} (2u)^{\frac{1}{\eta_m+1}} - 1 & \text{if } u < 0.5 \\ 1 - [2(1-u)]^{\frac{1}{\eta_m+1}} & \text{otherwise .} \end{cases} \tag{8}$$

4.4 The Evolutionary Engine

There are multiple state-of-the-art algorithms for MO optimization [10]. In this work an approach based on the NSGA-II algorithm, [11], is utilized. The main characteristics of the evolutionary engine are the use of ranking based generational elitism and the incorporation of crowding distance measures within each rank. At each iteration the Parent and Child population of the previous are joined into a temporary population. This composite population is ranked according to Pareto dominance. This iterative procedure obtains the Pareto front of the population and assigns to each of its individuals a rank value of 1. Afterward, these individuals are eliminated from consideration and a new Pareto Front is calculated from the remaining population. The elements of such "local" Pareto Front are assigned a rank value of 2. This is repeated until each individual

has been ranked, increasing the rank value at each iteration. In this way, multiple "layers" or ranks of Pareto Fronts are calculated. After the ranking process has finished, all the elements of the first rank (the Pareto Front of the whole population) are assigned the same fitness. In order to promote diversity among the solutions in each rank, crowding penalties are imposed to the closely located individuals. After such penalization has been enforced on the fitness values, the process is repeated for the next rank. Special care is taken to insure that the lowest fitness found in a given rank is greater that any of the individuals on the next. This gives selective pressure to the individual with lower rank (e.g. closer to the Pareto Front of the whole population). The next population of Parents is determined by an elitist strategy that selects elements in consideration of their rank. This new Parent population generates a new Child population and the process is repeated. Figure 1 shows a high level description of our approach.

5 Experimental Results

Experimentation was carried out for the simulation of a complex three dimensional object under observation by a manipulator robot. The goal is to obtain a photogrammetric network that is optimal in terms of reconstruction accuracy and manipulator motion. It is reasonable to assume that the most *efficient* configuration, in terms of motion alone, will be one where the robot takes images from the same viewpoint. On the other hand, an optimal network in terms of precision exclusively, will be a well distributed set of viewpoints. Using a population of N=100 individuals the evolutionary algorithm was executed for a period of $t_{max} = 100$ generations. The crossover probability was $P_x = 0.8$ with a spread factor $\eta_x = 2$. The mutation probability was $P_m = 0.012$ with $\eta_m = 2$. The execution time was of 95 seconds on a Pentium 4 at 2.4 GHz. The resulting Pareto Front is illustrated in Figure 2. The horizontal axis reflects the magnitude of the 3D reconstruction uncertainty. The vertical axis illustrates the total distance traveled by the manipulator. A single point in this plot represents the function values corresponding to a given sensing specification. The asymptotic behavior on both axes, as well the convex shape of the Pareto Front, support the assumption that our sensor planning problem is indeed MO. Some of the corresponding network configurations to the set of non dominated solutions are presented in Figure 3. In the absence of prior knowledge about the real Pareto Front, it is difficult to derive conclusions on the quality of the obtained non dominated solutions in terms of convergence. However, notice how the solutions with better efficiency, Solution F in Figure 3, approach the aforementioned ideal configuration in terms of motion. On the opposite extreme of the obtained Pareto Front, Solution A in Figure 3, the most accurate configuration is similar to an optimal imaging geometry presented in [7].

6 Conclusions and Future Work

A novel MO approach to sensor planning has been presented. An EC methodology based on the NSGA-II algorithm was described, along with experimentation

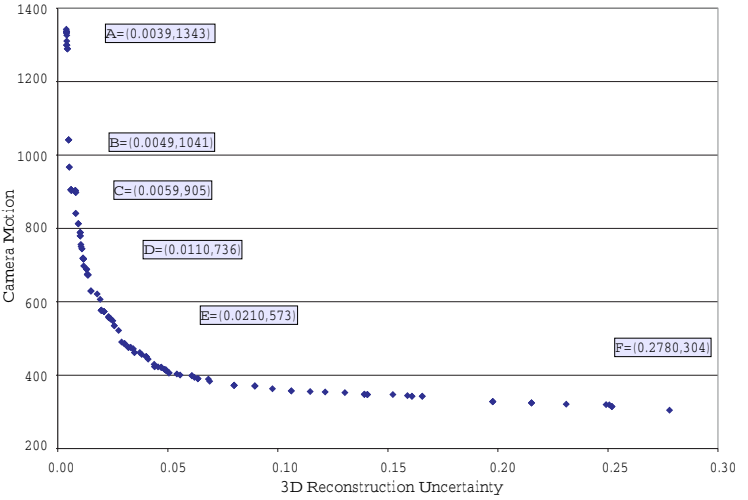

Fig. 2. Pareto Front for the case of Accuracy vs. Motion. Indicated optimal configurations are presented in Figure 3.

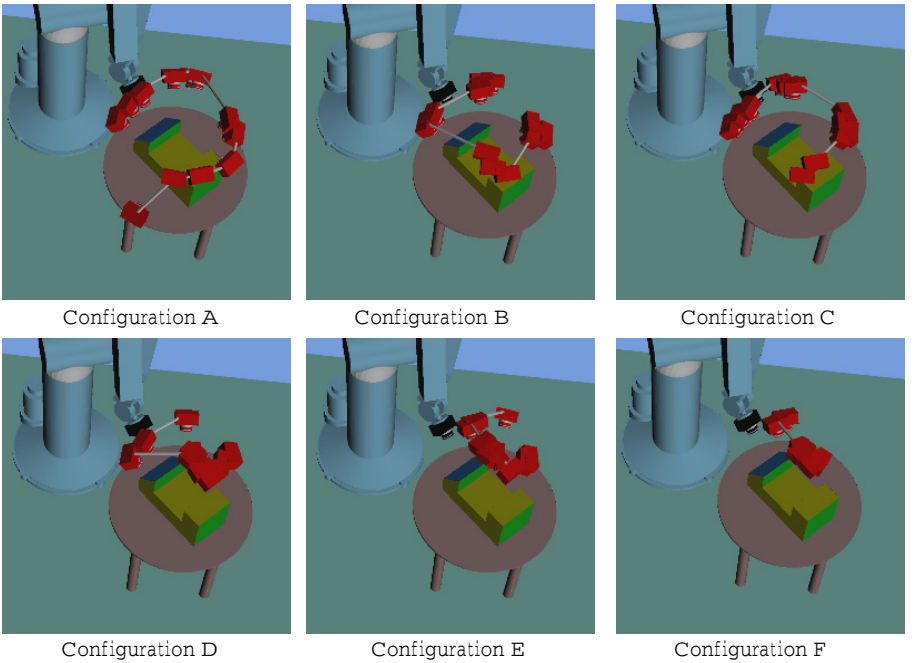

Fig. 3. Different optimal configurations. The correspondence of the presented solutions with the Pareto Front is illustrated on Figure 2.

on a simulation environment. The results comply with the assumption of conflicting objectives within the sensing task specification. Moreover, the optimal solutions found by the evolutionary algorithm form a convex Pareto Front with asymptotic behavior in each axis. The developed system is an efficient and expandable tool for the study of the MO sensor planning problem. Future research includes the implementation of our approach in a real world environment, the use of alternative optimization engines and operators, relaxing the viewpoint placement constraints and studying more complex scenarios. Particularly, the encapsulation of the crossover and mutation operators into the approach proposed in [12] is expected to improve the performance of our system.

Acknowledgments. Research funded by Contract 35267-A from CONACyT and under the LAFMI Project. First author supported by scholarship 142987 from CONACyT.

References

1. Tarabanis K.A., Allen P.K., Tsai R.Y. 1995. "A Survey of Sensor Planning in Computer Vision". IEEE Trans. on Robotics and Automation. 11(1):86-104 p.
2. Newman T., Jain A. 1995. "A Survey of Automated Visual Inspection", Computer Vision and Image Understanding 61(2):231-262.
3. Mason S. 1997. "Heuristic Reasoning Strategy for Automated Sensor Placement. Photogrammetric Engineering & Remote Sensing, 63(9):1093-1102 p.
4. Sakane S, Niepold R, Sato T, Shirai T. 1992. "Illumination Setup Planning for a Hand-eye System Based on an Environmental Model". Advanced Robotics. 6(4).
5. Tarabanis K., Tsai R., Allen P.K. 1995. "The MVP Sensor Planning System for Robotic Vision Tasks", IEEE Trans. on Robotics and Automation, 11(1).
6. Ikeuchi K., Robert J.C. 1991. "Modeling sensors detectability with the VANTAGE geometric/sensor modeler". IEEE Trans. on Robotics and Automation, Vol. 7.
7. Olague G. 2002. "Automated Photogrammetric Network Design Using Genetic Algorithms". Photogrammetric Engineering & Remote Sensing, 68(5):423-431. Awarded "2003 First Honorable Mention for the Talbert Abrams Award", by ASPRS.
8. Dunn, E. and Olague, G. 2003. "Evolutionary Computation for Sensor Planning: The Task Distribution Plan".EURASIP Journal on Applied Signal Processing.
9. Olague G. Mohr R. 2002. "Optimal Camera Placement for Accurate Reconstruction". Pattern Recognition, 35(4):927-944 p.
10. Coello C., Van Veldhuizen D., Lamont G. 2002. "Evolutionary Algorithms for Solving Multi-Objective Problems". Kluwer Academic Publishers. New York, NY.
11. Deb K., Agrawal S., Pratab A., Meyarivan T., 2000. "A Fast Elitist Non-Dominated Sorting Algorithm for Multi Objective Optimization: NSGA-II". Proceedings of PPSN VI, Paris France. Springer. Lecture Notes in Computer Science No. 1917.
12. Olague, G., Hernández, B. and Dunn, E., 2003. "Accurate L-corner Measurement using USEF Functions and Evolutionary Algorithms". 5th European Workshop on Evolutionary Computation in Image Analysis and Signal Processing. Lecture Notes in Computer Science 2611. Springer-Verlag.