Automated Photogrammetric Network Design Using the Parisian Approach

Enrique Dunn[1], Gustavo Olague[1], and Evelyne Lutton[2]

[1] Centro de Investigación Científica y Educación Superior de Ensenada, División de Física Aplicada, EvoVisión Laboratory
{edunn, olague}@cicese.mx
[2] INRIA - COMPLEX Team, Domaine de Voluceau BP 105 78153 Le Chesnay Cedex - France
Evelyne.Lutton@inria.fr

Abstract. We present a novel camera network design methodology based on the *Parisian* approach to evolutionary computation. The problem is partitioned into a set of homogeneous elements, whose individual contribution to the problem solution can be evaluated separately. These elements are allocated in a population with the goal of creating a single solution by a process of aggregation. Thus, the goal of the evolutionary process is to generate individuals that jointly form better solutions. Under the proposed paradigm, aspects such as problem decomposition and representation, as well as local and global fitness integration need to be addressed. Experimental results illustrate significant improvements, in terms of solution quality and computational cost, when compared to canonical evolutionary approaches.

1 Introduction

Automatic camera placement for artificial perception tasks consists on deciding the position of a set of sensors with respect to an observed scene. Depending on the selected task, the resulting problem offers an intricate combination of interactions between the sensor physical constraints, the mathematical modeling of the problem, as well as the numerical methods used to solve it [1], [2]. Accurate 3D reconstruction is a particularly difficult problem that needs to be automated. The complexity is mainly due to the stochastic nature of the uncertainty assessment process which requires multiple redundant image measurements [3]. Indeed, a difficult numerical adjustment problem for 3D reconstruction arises along with a highly discontinuous design search space for imaging geometry. The Bundle Method for optical triangulation is a fundamental aspect involved in the attainment of precise mensuration in photogrammetric projects. Indeed, the lack of a widespread utilization outside this community can be attributed to its expensive computational requirements and inherent design complexity. However, the development of an effective camera network configuration should also be based on a rigorous photogrammetric approach. This work presents the ongoing

F. Rothlauf et al. (Eds.): EvoWorkshops 2005, LNCS 3449, pp. 356–365, 2005.

development of the EPOCA [4],[5] sensor planning system and implements an evolutionary computation methodology based on the Parisian approach. This is done in order to efficiently search the space of possible camera configurations while maintaining high qualitative solutions of the photogrammetric adjustment process.

Photogrammetric Network Design is an active research field in photogrammetry, see [6], where recent works have provided important insights into the problem of determining an optimal imaging geometry. Mason [7] proposed an expert system approach based on the theory of generic networks in order to automate the viewpoint selection process. Thus, the decision making is carried out by heuristic means utilizing extensive expert prior knowledge. On the other hand, the work of Olague [8] uses an evolutionary computation approach, developing a criterion based on the error propagation phenomena. In this way, the design search space is explored by an stochastic meta-heuristic that yields human competitive results.

2 Problem Statement: Photogrammetric Network Design

Accuracy assessment of visual 3D reconstruction consists on attaining some characterization of the uncertainty of our results. The design of a photogrammetric network is the process of determining an imaging geometry that allows accurate 3D reconstruction. Rigorous photogrammetric approaches toward optical triangulation are based on the bundle adjustment method [9], which simultaneously refines scene structure and viewing parameters for multi-station camera networks. Under this nonlinear optimization procedure, the image forming process is described by separate functional and stochastic models. The functional model is based on the collinearity equations given by $s(\mathbf{p}-\mathbf{c}^p) = \mathbf{R}(\mathbf{P}-\mathbf{C}^o)$, where s is a scale factor, $\mathbf{p}$=$(x, y, -f)$ is the projection of an object feature into the image, $\mathbf{c}^p = (x^p, y^p, 0)$ is the principal point of the camera , $\mathbf{P} = (X, Y, Z)$ represents the position of the object feature, $\mathbf{C}^o = (X^o, Y^o, Z^o)$ denotes the optical center of the camera, while $\mathbf{R}$ is a rotation matrix expressing its orientation.

This formulation is readily extensible to multiple features across several images. For multiple observations a system of the form $\mathbf{l} = f(\mathbf{x})$ is obtained after rearranging and linearizing the collinearity equations, where $\mathbf{l} = (x_i, y_i)$ are the observations and $\mathbf{x}$ the viewing and scene parameters. Introducing a measurement error vector $\mathbf{e}$ we obtain a functional model of the form $\mathbf{l} - \mathbf{e} = \mathbf{Ax}$.

The design matrix $\mathbf{A}$ is of dimension $n \times u$, where n is the number of observations and u the number of unknown parameters. Assuming the expectancy $E(\mathbf{e}) = 0$ and the dispersion operator $D(\mathbf{e}) = E(\mathbf{ee}^t) = \sigma_0^2\mathbf{W}^{-1}$, where $\mathbf{W}$ is the "weight coefficient" matrix of observations, we obtain the corresponding stochastic model: $E(\mathbf{l}) = \mathbf{Ax}\Sigma_{ll} = \Sigma_{ee} = \sigma_0^2\mathbf{W}^{-1}$. Here Σ is the covariance operator and σ_0^2 the variance factor. The estimation of $\mathbf{x}$ and σ_0^2 can be performed by least squares adjustment in the following form

$$\mathbf{x} = (\mathbf{A^TWA})^{-1}\mathbf{A^TWl} = \mathbf{QA^TWl},$$

$$\mathbf{v} = \mathbf{A}\mathbf{x} - \mathbf{l} \qquad \sigma_0^2 = \frac{\mathbf{v^t W v}}{r}$$

where r is the number of redundant observations, $\mathbf{v}$ is the vector of residuals after least squares adjustment and $\mathbf{Q}$ is the corresponding cofactor matrix. Additionally, the covariance of the parameters is given by $\Sigma_{xx} = \sigma_0^2\mathbf{Q}$. The vector of parameters can be separated in the form $\mathbf{x} =< \mathbf{x_1}, \mathbf{x_2} >$, where $\mathbf{x_1}$ contains the viewing parameters while $\mathbf{x_2}$ expresses the scene structure correction parameters. Thus, we obtain a system of the form

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{A_1^T W A_1} & \mathbf{A_1^T W A_2} \\ \mathbf{A_2^T W A_1} & \mathbf{A_2^T W A_2} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{A^T W l} \\ \mathbf{A^T W l} \end{pmatrix}$$

Accordingly, the cofactor matrix $\mathbf{Q}$ can be written as

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_{1,2} \\ \mathbf{Q}_{2,1} & \mathbf{Q}_2 \end{pmatrix}$$

The matrix $\mathbf{Q}_2$ describes the covariance structure of scene coordinate corrections. Hence, an optimal form of this matrix is sought in order to obtain accurate scene reconstruction. The criterion we selected for minimization is the average variance along the covariance matrix, see [8],

$$f_1(\mathbf{x_1}, \mathbf{x_2}) = \sigma^2_{\mathbf{x}_2} = \frac{\sigma_0^2 \ trace(\mathbf{Q}_2)}{3n}. \tag{1}$$

3 The Parisian Approach

The Parisian Approach differs from typical approaches to evolutionary computation in the sense that a single individual in the population represents only a part of the problem solution. In this respect, it is similar to the Michigan approach developed for Classifier Systems. Hence, an aggregation of multiple individuals must be considered in order to obtain a solution for the problem being studied. Thus, the evolution of the whole population is favored instead of the emergence of only a single dominant solution. The motivation for such an approach is to make an efficient use of the genetic search process. This can be achieved from two different perspectives. First, the algorithm discards less computational effort at the end of execution, while considering more than a single best individual as output. Second, the computational expense of the fitness function evaluation is considerably reduced for a single individual.

Successful examples of such an approach can be found in the image analysis and signal processing literature. The *Fly Algorithm* developed by Louchet et al. [10] is a real-time pattern recognition tool used in stereo vision systems. In such a work, the population is formed by individuals representing each a single 3D point. The evolutionary algorithm favors the positioning of each so called “fly” to a surface point in the observed scene using insightful problem modeling. The work of Raynal et al.[11] incorporates the Parisian approach to the solution

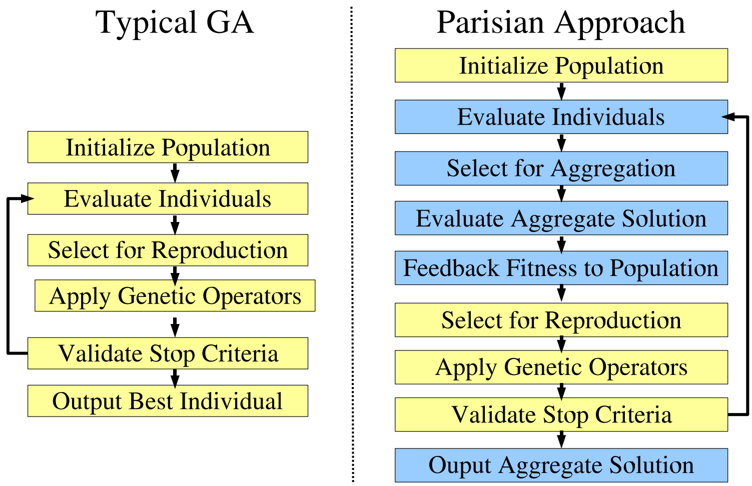

Fig. 1. Outline of our implementation of the *Parisian Sensor Planning* approach. Fitness evaluation is modified in order to consider the local and global contribution of an individual

of the inverse problem for Iterated Function Systems (IFS). In this instance a Genetic Programming methodology was adopted and experimentation on 2D images presented.

Many of the canonical aspects of evolutionary algorithms are retained under the Parisian approach, see Figure 1, allowing for a great flexibility in its deployment. However, the applicability of this paradigm is restricted to problems where the solution can be decomposed into homogeneous elements or components, whose individual contribution to the complete solution can be evaluated. Therefore, each implementation is necessarily application dependent, where the design of a suitable problem decomposition is determinant factor. Thus, the following implementation issues have been identified:

- **Partial Encoding.** The genetic representation used for a single individual encodes a partial solution.
- **Local Fitness.** A meaningful merit function must be designed for each partial solution. In this way, the worthiness of a single individual can be evaluated in order to estimate the potential contribution to an aggregate solution.
- **Global Fitness.** A method for the aggregation of multiple partial solutions must be determined. In turn, a problem defined fitness function can be evaluated from this complete solution. However, the worthiness of this composite solution should be reflected on each partial solution.
- **Evolutionary Engine.** The evolution of the complete population should promote the emergence of better aggregate solutions. The evolutionary engine requires a scheme for combining local and global fitness values. Also,

it requires a diversity preserving mechanism in order to maintain a set of complementary partial solutions.

4 Parisian Approach to Camera Network Design

Camera placement can be viewed as a geometric design problem where the control variables are the spatial positioning and orientation of a finite set of cameras. In order to state such design problem in optimization terms the criterion expressed in section 2 is adopted. However, due to the sensor characteristics and mathematical modeling of the problem a strongly constrained optimization problem emerges. In this section we will discuss the different implementation issues involved in our incorporation of the Parisian approach into the sensor planning problem.

4.1 Problem Partitioning and Representation

A viewing sphere model for camera placement is adopted in order to reduce the dimensionality of the search space. Therefore, each camera position is defined by its polar coordinates $[\alpha_i, \beta_i]$. A network of M cameras is represented by a real valued vector

$$\boldsymbol{\Psi} \in \mathbb{R}^{2M} \quad where \quad \alpha_i = \Psi_{2i-1}, \beta_i = \Psi_{2i} \quad for \quad i = 1, \ldots, M. \tag{2}$$

Our design problem allows the decomposition into individual elements since the complete camera network is formed by a set of homogeneous components. Nevertheless, a decision on the level of *granularity* of our decomposition is crucial. Here we have the choice of an individual representing a single camera or a camera subnetwork (i.e. a set of cameras). We have decided for the latter option since such an individual can be meaningfully evaluated in terms of its imaging geometry contribution to 3D reconstruction. Hence, each individual in the population represents a fixed size subnetwork of N cameras, denoted by a vector of the form

$$\boldsymbol{\psi}^j \in \mathbb{R}^{2N} \quad where \quad \alpha_i = \boldsymbol{\psi}^j_{2i-1}, \beta_i = \boldsymbol{\psi}^j_{2i} \quad for \quad i = 1, \ldots, N, \tag{3}$$

where j is defined as the subnetwork population index. Accordingly, a complete camera network specification is given by the aggregation of J subnetworks

$$\boldsymbol{\Psi} \in \mathbb{R}^{2M} = \bigcup_{j=1}^{J} \boldsymbol{\psi}^j, \quad where \quad M = J \times N \tag{4}$$

4.2 Local Fitness Evaluation

Section 2 presented a photogrammetric approach for estimating the variance of 3D point reconstruction using redundant measurements, see Eq. (1). Such methodology is generally applied to the complete measured object considering all cameras concurrently. Since in our representation we are working with camera subnetworks, it is unlikely that any single individual successfully captures the complete 3D object denoted by the whole set of 3D points $\mathbf{P}$. Hence, the object is also partitioned into R disjoint regions or subsets of points, in such a way that $\mathbf{P} = \bigcup_{i=1}^{R} P_i$. The visibility of a camera subnetwork $\boldsymbol{\psi}^j$ is limited to a subset of the whole object, expressed by $\mathbf{V}(\boldsymbol{\psi}^j) \subseteq \mathbf{P}$. Therefore, we define the field of view constraint in the form

$$C_{fov}(\boldsymbol{\psi}^j, P_i) = \begin{cases} 1 & \text{if } P_i \subset \mathbf{V}(\boldsymbol{\psi}^j) \\ 0 & \text{otherwise .} \end{cases}$$

Thus, the local fitness evaluation uses the idea of decomposing the problem in subnetworks which provide greater object coverage with higher precision in order to attain higher fitness values. The reconstruction uncertainty for each set P_i is evaluated for a single individual accordingly to Equation (1), discarding the portions of the object not sensed by such a subnetwork. Thus, we have

$$f_{loc}(\boldsymbol{\psi}^j) = \sum_{i=1}^{R} \frac{1}{f_1(\boldsymbol{\psi}^j, P_i)} \quad \forall P_i : C_{fov}(\boldsymbol{\psi}^j, P_i) = 1. \tag{5}$$

4.3 Global Fitness Evaluation

Once the local fitness of each individual has been evaluated, a process of aggregation is needed to obtain a solution to our camera network design problem. In order to achieve this, a *selection* of a group of individuals from the population must be made. The selection should be based on the merit of each individual fitness and can be realized by deterministic (i.e. selecting the top J individuals in the population) or stochastic (i.e. roulette, tournament) means. In this way, at each generation t an aggregate solution $\boldsymbol{\Psi}(t)$ has been formed for global fitness evaluation. This global evaluation uses the same criterion as in the local fitness evaluation. Therefore, we obtain:

$$f_{global}(\boldsymbol{\Psi}(t)) = \sum_{i=1}^{R} \frac{1}{f_1(\boldsymbol{\Psi}(t), P_i)} \quad \forall P_i : C_{fov}(\boldsymbol{\Psi}(t), P_i) = 1. \tag{6}$$

Such value describes the aptitude of the aggregate solution obtained at generation t. Obviously, the goal of the algorithm is to foster the improvement of this global fitness along the course of successive generations. However, another purpose of this evaluation is to be able to reflect on the population the effect of the evolutionary process. The individuals that form part of the aggregate solution will be rewarded or punished based on its global fitness. Also, based on the complete solution characteristics, promising individuals not selected should be compensated so they might contribute in latter stages of the evolutionary process.

A valid solution to the network design problem is one that reconstructs accurately the complete object. This requires addressing the aspects of constraint satisfaction and function optimization. We shall now describe how we use global fitness evaluation to deal concurrently with both of these issues.

Function optimization will be addressed first. In order to reflect the quality of an aggregate solution $\boldsymbol{\Psi}(t)$ on each of the individuals $\boldsymbol{\psi}^j$ that compose it, we use the ratio of improvement in global fitness among successive generations. The magnitude of the adjustment of an individual's local fitness is proportional to this ratio as follows

$$g_1(\boldsymbol{\psi}^j) = f_{loc}(\boldsymbol{\psi}^j)\left[\frac{f_{global}(\boldsymbol{\Psi}(t))}{f_{global}(\boldsymbol{\Psi}(t-1))} - 1\right] \qquad \forall \boldsymbol{\psi}^j \in \boldsymbol{\Psi}(t). \tag{7}$$

Now we shall consider constraint satisfaction. It is very likely that each individual subnetwork will only cover part of the object. It is also possible that a given aggregation of individuals will not provide full object coverage. In this respect, when a particular aggregate solution $\boldsymbol{\Psi}(t)$ does not cover some object region P_i (e.g. $C_{fov}(\boldsymbol{\Psi}(t), P_i) = 0$), it would be desirable to enhance the fitness value of those individuals on the population that indeed cover such region. The amount of enhancement of those individuals shall be proportional to their difference in fitness with respect to the best individual in the population. Hence, we have

$$g_2(\boldsymbol{\psi}^j) = f_{loc}(\boldsymbol{\psi}^{best}) - f_{loc}(\boldsymbol{\psi}^j) \qquad \forall \boldsymbol{\psi}^j : \mathbf{V}(\boldsymbol{\psi}^j) \bigcap \overline{\mathbf{V}(\boldsymbol{\Psi}(t))} \neq \emptyset. \tag{8}$$

Note that this value is only calculated for those individuals that cover an object region not sensed by the aggregate solution formed at that generation.

In this way, the global fitness is "fedback" to the general population in the following manner:

$$F(\boldsymbol{\psi}^j) = \begin{cases} f_{loc}(\boldsymbol{\psi}^j) + \lambda_1 g_1(\boldsymbol{\psi}^j) & \text{if } \boldsymbol{\psi}^j \in \boldsymbol{\Psi}(t) \\ f_{loc}(\boldsymbol{\psi}^j) + \lambda_2 g_2(\boldsymbol{\psi}^j) & \text{if } \mathbf{V}(\boldsymbol{\psi}^j) \bigcap \overline{\mathbf{V}(\boldsymbol{\Psi}(t))} \neq \emptyset \\ f_{loc}(\boldsymbol{\psi}^j) & \text{otherwise .} \end{cases}$$

Here, λ_1 and λ_2 are user defined parameters that reflect the relative importance given to each of the aspects involved in the global fitness evaluation.

5 Experimental Results

The reconstruction of a complex 3D object is considered in our experimentation. The goal is to determine a viewing configuration that will offer optimal results in terms of reconstruction accuracy. Here, we shall consider the design of a fixed size camera network of $M = 9$ stations. According to our approach, the level of granularity of our problem decomposition needs to be established. For these series of experiments we will use camera subnetworks of $N = 3$ cameras. In this way, each of the individuals in the population will consist of a vector $\boldsymbol{\psi} \in \mathbb{R}^6$. Hence, a total of $J = 3$ subnetworks will need to be aggregated in order to form a complete solution to our network design problem. The convex polyhedral

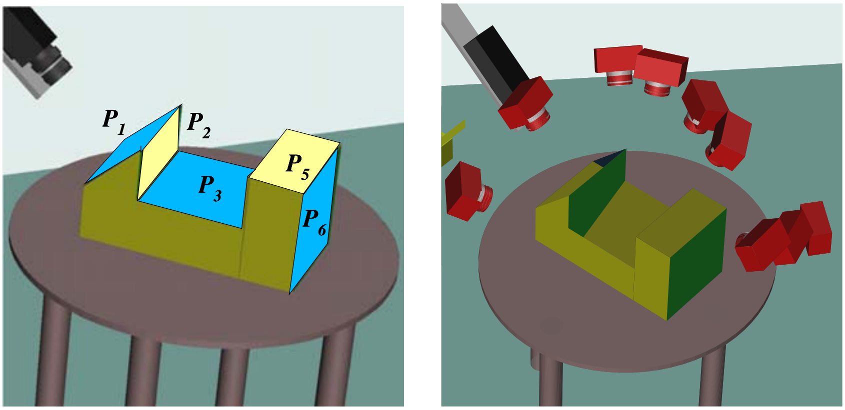

Fig. 2. The 3D object under observation. The concave object is partitioned into different regions in order to facilitate the fitness evaluation of sub-networks of small size. A photogrammetric network formed by 9 cameras is illustrated on the right

object under study, depicted in Figure 2, is partitioned into $R = 6$ regions. Elitist selection of individuals for solution aggregation is based on their fitness value. Finally, the user defined valued λ_1 and λ_2 are set to $\lambda_1 = \lambda_2 = 1.0$.

For all our experiments, SBX-crossover is utilized with a probability $P_c = 0.95$ along with polynomial mutation subject to probability $P_m = 0.05$. We have used tournament selection for reproduction under generational replacement. Alongside of our methodology, the same global fitness function was optimized by a typical genetic algorithm (e.g. each individual encodes a complete solution). This was done in order to have some reference point in the assessment of our proposed methodology. Both evolutionary algorithms were executed for 100 generations, using a population of 30 individuals.

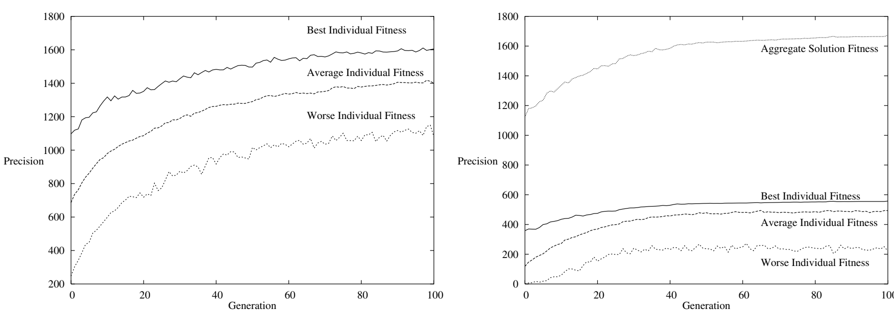

Fig. 3. Performance Comparison. On the left, the population evolution of a typical genetic algorithm is depicted. On the right, higher fitness values are consistently attained by the aggregate solutions of our proposed methodology. Plotted values reflect the averages over 20 executions

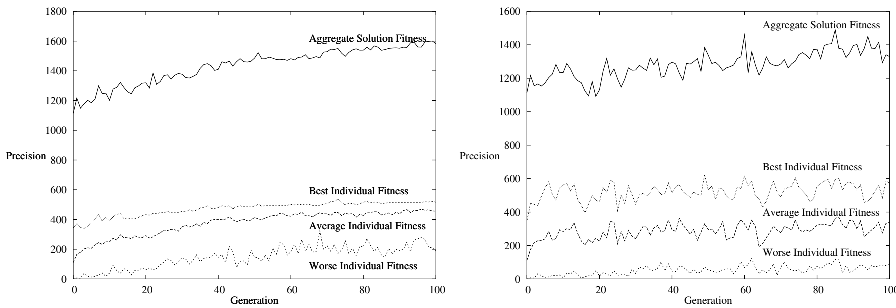

Fig. 4. Dependence on parameters λ_1, λ_2.The plot on the left corresponds to an execution with mixing values $[\lambda_1 = 0.8, \lambda_2 = 0.2]$. Performance is slightly deteriorated.The plot on the right represents an execution with values $[\lambda_1 = 0, \lambda_2 = 1.5]$. Note the almost random algorithm performance

Figure 3 plots population performance measures (best, mean, worse fitness) for a canonical GA on the left and also for our Parisian approach on the right. While these measures are descriptive of the dynamics of our population, the importance is on the *aggregate solution fitness* measure. In this respect, our approach slightly outperforms a canonical methodology in terms of solution quality. However, these results are made more relevant when considering the computational cost involved in fitness evaluation. For our studied object, evaluation based on the bundle adjustment of a complete network of 9 cameras is over 15 times more costly than that of a 3 camera subnetwork. Accordingly, by virtue of our problem decomposition, the total **execution time of the algorithm is reduced 10 times**. Clearly, a significant benefit in performance has been achieved.

The choice of mixing values λ_1, λ_2 is an important aspect in the performance of the algorithm, as they determine the magnitude of the global fitness adjustment given to each individual. In order to exemplify this, we have carried out different experiments varying the ratio and magnitude of these values. Experiments show a fairly robust behavior for similarly scaled values under 1.0. In general, performance deteriorates as the magnitude and the ratio among parameters increases. The right plot of Figure 4 illustrates the scenario where constraint satisfaction is completely predominant over function optimization. As a result, the fitness value of aggregate solutions is decreased by the inclusion of weaker subnetworks that are unreasonably enhanced by the global fitness evaluations.

6 Conclusions and Future Research

The Parisian approach to evolutionary computation offers an efficient way to address the design of photogrammetric networks. Experimental results illustrate its favorable performance against canonical evolutionary algorithms applied to our problem. In fact, solution quality is slightly improved with a 10 times reduction in computational effort. Aspects crucial to our methodology such as problem decomposition and representation, as well as the integration of local and global fitness evaluation have been discussed. However, further characterization of our algorithm behavior still is needed. Particularly, aspects like determining a suitable problem decomposition granularity and population size, the assignment of global fitness mixing values λ_1, λ_2 and the effect of diversity preservation techniques on the evolutionary process need to be addressed.

Acknowledgments

This research was funded by CONACyT and INRIA through the LAFMI project 634-212. First author supported by scholarship 142987 from CONACyT.

References

1. Olague, G.: A Comprehensive Survey of Sensor Planning for Robot Vision. Available as CICESE Research Report No. 25259.
2. Wong C., Kamel M.: Comparing Viewpoint Evaluation Functions for Model-Based Inspectional Coverage. 1st Canadian Conference on Computer and Robot Vision (CRV'04). May 2004. pp. 287-294.
3. Firoozfam P., Negahdaripour S.: Theoretical Accuracy Analysis of N-Ocular Vision Systems for Scene Reconstruction, Motion Estimation, and Positioning. 2nd International Symposium on 3D Data Processing, Visualization, and Transmission, (3DPVT'04). September 2004. pp. 888-895
4. Dunn, E. and Olague, G.: "Evolutionary Computation for Sensor Planning: The Task Distribution Plan". EURASIP Journal on Applied Signal Processing 2003:8,748-756.
5. Olague G. Mohr R.: "Optimal Camera Placement for Accurate Reconstruction". Pattern Recognition, 35(4):927-944 p.
6. Saadatseresht M., Fraser C., Samadzadegan F., Azizi A.: Visibility Analysis In Vision Metrology Network Design. The Photogrammetric Record 19(107):219-236. September 2004.
7. Mason S.: "Heuristic Reasoning Strategy for Automated Sensor Placement. Photogrammetric Engineering & Remote Sensing, 63(9):1093-1102 p.
8. Olague G.: "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.
9. McGlone C. (ed.): Manual of Photogrammetry. Published by American Society for Photogrammetry and Remote Sensing. Bethesda, Maryland 20814. 1151pp.
10. Louchet J., Guyon M., Lesot M. and Boumaza A.: Dynamic Flies: a new pattern recognition tool applied to stereo sequence processing. In Pattern Recognition Letters, No. 23 pp. 335-345. (2002)
11. Raynal F., Collet P., Lutton E. and Schoenauer M.: Polar IFS + Parisian Genetic Programming = Efficient IFS Inverse Problem Solving. In Genetic Programming and Evolvable Machines Journal, Volume 1, Issue 4, pp. 339-361. (2000)