INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Optimal Camera Placement for Accurate Reconstruction

Gustavo OLAGUE, Roger MOHR

N° 3338

Janvier 1998

THÈME 3

Rapport de recherche

ISSN 0249-6399

Optimal Camera Placement for Accurate Reconstruction

Gustavo OLAGUE, Roger MOHR

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet MOVI

Rapport de recherche n° 3338 — Janvier 1998 — 34 pages

Abstract: 3D measurements can be recovered from several views by triangulation. This paper deals with the problem of where to place the cameras in order to obtain a minimal error in the 3D measurements, also called *camera network design* in photogrammetry. We pose the problem in terms of an optimization design, dividing it into two main components: 1) an analytical part dedicated to the analysis of error propagation from which a criterion is derived, 2) a global optimization process to minimizes this criterion. In this way, the approach consists of an uncertainty analysis applied to the reconstruction process from which a covariance matrix is computed. This matrix represents the uncertainty of the detection from which the criterion is derived. Moreover, the optimization has discontinuities mainly due to the unobservability of points, which leads to a combinatorial optimization process. These aspects are solved using a multicellular genetic algorithm. Experimental results are provided to illustrate the effectiveness and efficiency of the solution.

Keywords: Camera network design, uncertainty analysis, global optimization, covariance matrix, genetic algorithms.

(Résumé : tsvp)

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN (France)
Téléphone : 04 76 61 52 00 - International: +33 4 76 61 52 00
Télécopie : 04 76 61 52 52 - International: +33 4 76 61 52 52

Placement Optimal de Caméras pour une reconstruction précise

Résumé : Les mesures tridimensionelles peuvent être obtenues à partir de plusieurs images par la méthode de triangulation. Cet article évoque le problème du placement des caméras de façon à obtenir une erreur minimale lors des mesures tridimensionelles. En photogrammétrie, on parlera du concept du réseau de caméras. Nous poserons le problème en termes d'optimisation et nous le diviserons en deux parties: 1) Une partie analytique dédiée à l'analyse de l'erreur de propagation d'où découlera un critère. 2) Un processus global d'optimisation minimisera ce critère. De ce côté-là, l'approche consiste en une analyse d'incertitude appliquée au processus de reconstruction d'où une matrice de covariance sera calculée. Cette matrice représente l'incertitude de la detection pour lequel le critère est dérivé. Par ailleurs, l'optimisation a des aspects discontinus essentiellement dû à l'inobservabilité des points. Ce facteur va nous amener à utiliser un processus d'optimisation combinatoire que nous avons résolu en utilisant un algorithme génétique multicellulaire. Des resultats expérimentaux sont inclus pour illustrer l'efficacité et la rapidité de la solution.

Mots-clés: Concept de réseau de caméras, analyse d'incertitude, optimisation globale, matrice de covariance, algoritmes génétiques.

Contents

List of Figures

List of Tables

1 Introduction

Visual measurements can be made significantly more accurate by selecting a well designed convergent camera configuration and a suitable mathematical model of the process. As an example see the work of Fraser [Fra92] in which he reports highly accurate measurements employing a strong (hand-designed) photogrammetric network using 36 photographs taken by 18 well disposed cameras with high convergence angles. Another example is the work of Beyer [Bey92] in which he reports accuracies of 1/100th of a pixel using CCD cameras. Despite achievements like these, and due to the complexity of designing the convergent multistation networks, the photogrammetric measurement technique has rarely been applied by other than experienced photogrammetrists.

The problem of automating the camera network design process in order to attain highly accurate measurements has received comparatively little attention given its practical importance. This paper deals with the problem of determining the attitude of several cameras observing some targets placed on a 3D object with the goal of achieving highly accurate measurements. The input data would be the a priori approximatively known set of target points to be measured in order to obtain the best accuracy in the final measurements. Several additional constraints like the upper bound on the number of cameras and placement limitation due to the incidence angle must be taken into account in the exploration of the search space. Automating such a camera placement process reaches a high complexity due to the competing nature of the constraints and numerous design decisions that need to be made. The mathematical model of triangulation plays a role as it leads to the uncertainty model and therefore to the error criterion to be optimized. This paper presents a prototype system called EPOCA (an acronym for "Evolving POsitions of CAmeras") which we have implemented to test the validity of our approach.

The paper is divided as follows: firstly, we pose the problem isolating several key issues and review the literature related to sensor planning. We describe our approach into two main sections. The first section is divided into three subsections: the reconstruction process, 3D error estimation and image error estimation. All are used in the development of a criterion. The second section presents our method of camera placement implemented as a multicellular genetic algorithm which optimizes this criterion. Finally, we present some experimental results followed by a discussion.

1.1 Statement of the Problem

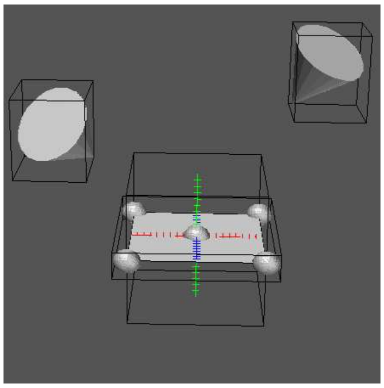


Figure 1: *Two convergent cameras. The cones illustrate the camera attitudes and the ellipsoids represent the uncertainty of the observed targets.*

The problem we would like to approach is the automation of the camera network design process in order to obtain accurate 3D measurements. We restrict ourselves to the problem where there are no limitation on the camera positions (the only exception is the incidence angle constraint.), and it is simplified to the case where the cameras remain at a fixed distance to the set of target points to be measured. Imagine that we would like to accurately measure some fiducial marks distributed over a planar surface, and that we would like to place the cameras to obtain a minimal error. Figure 1 shows two cameras observing five target points, represented by error ellipsoids describing their position uncertainty. Changing the camera attitudes (position and orientation) changes the orientation and size of the ellipsoids. The main question we would like to answer is given in Table 1. From these initial questions the choice of a **criterion** combined with an **optimization** process will be the key concepts.

- Where should we place the cameras in order to obtain the minimal 3D error?

  From this question several subproblems arise:

  1. How can we develop a good criterion to judge our configuration?
  2. What conditions are needed for our system to work?
  3. Which are the interrelated aspects involved in the development of the system?
  4. What would be a good method to optimize the placement of the camera?

Table 1: *Some initial questions.*

1.2 Related Work

Researchers in the computer vision and photogrammetric communities have recognized the need to automate the process of camera network design. The literature devoted to this problem is however limited due to different complex aspects that are involved. This section surveys the existing literature on automated single and multiple camera placement.

Single camera systems include the following. The HEAVEN system developed by Sakane et al. [SS91] is an example of a system that finds possible sensor positions using a *generate-and-test strategy*, for the inspection of an object tessellated by a sphere of a given radius. The object is unoccluded but an analysis of light source placement is incorporated. This study was extended to the VIO system developed by Sakane and Niepold [SNSS92]. Another example of the generate-and-test approach for camera and light positions is ICE [YHS95]. This places emphasis on the illumination planning component of the problem. For camera placement the main task constraint considered is edge visibility. The sensor is positioned to minimize the occlusion of selected feature edges. The evaluation of this criterion is based on an aspect graph representation of the object.

Other works have adopted a *synthesis approach*. Instead of taking a discrete approach, the idea is to model constraints as analytical functions. *The Automatic Sensor and Illuminator Positioning Work of SRI* carried out by Cowan et al. [CK88] optimizes the camera locations to satisfy several geometric constraints like: visibility, polyhedral objects in focus, within the sensor field of view, at sufficient resolution and incidence angle, and unoccluded by the object itself or other objects in the working environment. The MVP system developed by Tarabanis et al. [TTA91] determines an optimal sensor location and sensor parameters (focal length, focus setting and aperture) for viewing a set of surfaces and avoiding occlusion. MVP is currently being extended to work in dynamic environments [AAT]. The placement of illumination sources for inspecting edge features has also been investigated by Cowan et al. While all of these systems provide solutions to the task of viewing an object, they do not address the task of deriving accurate object dimensions. They are also inadequate for complex objects on which not all features will be visible from a *single* viewpoint.

Tarbox and Gottschlich [TG93] have recognised the need for *multistation solutions* to overcome object occlusion problems. They have implemented a solution in the IVIS system for an active triangulation sensor. Cowan has also experimented methods to place multiple sensors overcoming the occlusion problem associated with 3D objects. Fritsch and Crosilla [FC90] have investigated the potential of optimizing multi-station configurations using an analytical first order design (FOD) approach by iteratively shifting the sensor stations until the covariance matrix of the estimated object feature coordinates is better than a criterion matrix. However their approach is limited to the optimization of existing configurations. Critical decisions, for example the number of camera stations and their positioning for a good approximation to a strong configuration, are expected to be made by the user prior to the optimization. An infinite number of possible camera station poses exists just limited by the set of competing constraints. The complexity of camera placement belongs to the class of NP-complete[1] problems [Mas97]. Consequently, it can be concluded that a trial and error approach to FOD is unsatisfactory

[1]All known exact algorithms for NP-complete problems run, in the worst case, in a time exponential in the size of the input data. It is widely conjectured that this will not change in the future.

and not tractable. Finally, a work called CONSENS which follows the expert system approach and uses multiple cameras in combination with optical triangulation, was developed by Mason et al. [MG95]. It outlines a method of overcoming these limitations. The method is based on the theory of *generic networks*, which constitutes compiled expertise, describing an ideal configuration of four camera stations that can be employed to provide a strong imaging geometry for the class of planar network design problems. Complex objects are divided into planes, each one is evaluated through one of these networks and then connected with some extra cameras with the purpose to establish just one common datum.

Unlike all other approaches, our idea is to pose the problem in terms of a global optimization design which is capable of managing the problem using an adaptive strategy. It explores the solution space using both non continous optimization and combinatorial search.

2 The Approach

This section presents a method to approach the problem of optimal camera placement in order to obtain accurate positions of 3D target points. This approach can be divided into two main components. Firstly, we will develop an analytic uncertainty analysis based on error propagation phenomenon. This will allow us to express an error criterion to be minimized. Secondly, we will present an evolutionary optimization method similar to genetic algorithm [Gol89], which optimizes this criterion.

2.1 Three Dimensional Reconstruction

The problem involves a set of M cameras viewing P points placed over N surfaces. We want to know the best camera network in order to reduce the uncertainty of the reconstructed points. Given the camera matrices M, we can reconstruct the 3D points in a least square sense. We assume here that cameras have no distortion or, as usually in high accurate measurement, that distortion has been cancelled out after having been estimated. So, a camera can be considered as a system that performs a known linear projective transformation from the projective space $\mathbb{P}^3$ into the projective plane $\mathbb{P}^2$.

$$p_{ij} \sim M_i P_j \ . \tag{1}$$

This mapping is represented by a matrix $M_{3\times4}$

$$M = K(R \mid T) \ , \tag{2}$$

where K is the matrix of intrinsic parameters, R is a rotation matrix and T a translation vector. Equation (1) can be rewritten as

$$A(p_i, M_i)P = b(p_i, M_i) \ , \tag{3}$$

where A is a $2i \times 3$ matrix and b is a $2i \times 1$ vector. If matrix A^tA is invertible we can find the least squares solution

$$P = (A^tA)^{-1}A^tb \ , \tag{4}$$

which minimizes $\|AP - b\|^2$.

2.2 3D Error Estimation

Until now, we have studied the function to transform a point in space into an image point

$$P = f(p) \ , \tag{5}$$

given by Equation (4) which is going to be useful to develop an analysis of error propagation [Fau93]. The key of manipulating geometric uncertainty is to be able to transform the information or probability density function on a feature available in one form (image point) into another form of interest (point in space). This transformation of information can be grouped into a family of transformations which we approximate to the exact transformation by a first-order relation using Taylor series. Successive moments can be found by equating higher order terms; however, using higher order terms is neither viable nor desirable, as any computational simplicity will be lost. In this way, a linear approximation is to be used in which we assume a Gaussian distribution. Then the mean $E[P]$ and covariance ΛP are sufficient information to completely define the feature density function. All this is given by the following proposition (see [Fau93, chapter V]):

Proposition 1 *Given a random variable $p \in \mathbb{R}^n$, 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:*

$$\Lambda P = \frac{\partial f(E[p])}{\partial p} \Lambda p \frac{\partial f(E[p])}{\partial p}^t . \tag{6}$$

Therefore, ΛP is a symmetric positive definite matrix, which describes the bounds on $P = f(p)$ in the vicinity of $E[P] = f(E[p])$ given those of p in the vicinity of $E[p]$. This proposition lets us compute the uncertainty of the three-dimensional point knowing the uncertainty in the image points. Another model is needed to give an interpretation of matrix Λp, which describes bounds on the possible values of the coordinates of p. This implies the necessity of an image error model.

2.3 Image Error Estimation

In order to estimate the covariance matrix Λp of our 2D measured points, we will use Proposition 1. Therefore, we must have a function to relate image point errors taken from several photographs over different angles. The relation will be established using the cross-ratio [Moh93]:

$$f(k) = k(x_1, x_2, x_3, x_4) = \frac{(x_3 - x_1)(x_4 - x_2)}{(x_4 - x_1)(x_3 - x_2)} , \tag{7}$$

which is a projective invariant. In this way

$$\Lambda k = \partial f(k) \; \Lambda p \; \partial f(k)^t , \tag{8}$$

computes the uncertainty of the cross-ratio with respect to the uncertainty of the image points. It can be rewritten as

$$\sigma_p = \sigma_k \frac{\lambda}{||J_k||^2} , \tag{9}$$

where σ_k represents the standard deviation of the cross-ratio, λ is the average distance between the targets, and J_k is the Jacobian of k. We assume that the image errors are Gaussian and identical (similar targets) along their alignment direction, and constant during the experiment (similar distance of observation and viewing angle). We used a subpixel target detector, on the calibration grid of Figure 3, similar to one presented by Gruen [Gru85]. From hundred of retro-measures of four approximatively equidistant points along each direction, we compute the uncertainty in the image points σ_p from the measured uncertainty of the cross-ratio σ_k, using Equation (9). σ_k is a function of the observation angle α along their alignment direction. As the targets are circular and viewed as an ellipse, it can be assumed that the covariance of its two dimensional location can be considered as a diagonal matrix within the reference frame of the two axes of the ellipse, and that the associate variances are the one corresponding to the viewing angle in the two corresponding directions. The graphs, in Figure 2, show final results of our experiments along the x and y directions. These graphs can be approximated to a model using the Levenberg-Marquardt method [PFTV88] and the merit function[2]

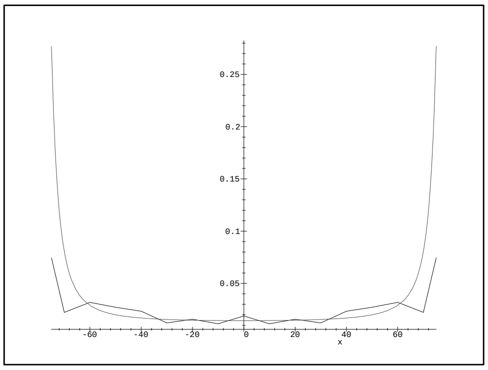
$$y = \beta(e^{(\frac{\alpha}{90-x})} + e^{(\frac{\alpha}{90+x})}) + \gamma \ , \tag{10}$$

yielding the best fit parameters: $\alpha = 79.74$, $\beta = 1.31 \times 10^{-3}$ and $\gamma = 8 \times 10^{-3}$ for both experiments. This model corresponds to the incidence angle constraint in camera network design. Notice that for an angle longer than 80 degrees, we have no more measurements and approaching this value leads to an infinite uncertainty. Λp model is useful to compute the covariance matrix ΛP of the 3D points.

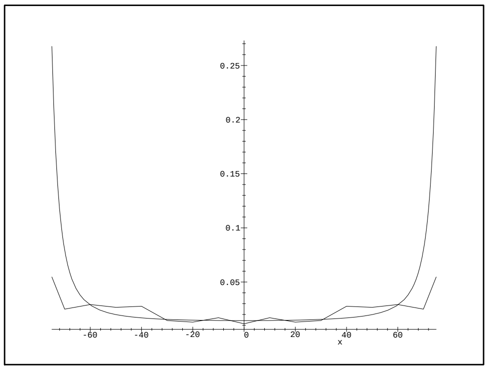
2.4 The Criteria

Once we have computed the covariance matrix ΛP, it is necessary to choose a criterion useful to the optimization process. In this sense, we need to select a metric to compare symmetric positive definite matrices. The *comparison* of covariance matrices is interpreted as the required standard deviation of function $P = f(p)$ to be better when it is computed with covariance matrix

[2]Note that these observations are better approximated to a hyperbolic function.



a) Best fit along x axis direction.



b) Best fit along y axis direction.

Graphs showing the best fit curve .

Figure 2: *The following graphs show the best fit curve along both axes. The X axis corresponds to the incidence angle constraint (in degrees) and the Y axis to the image error.*

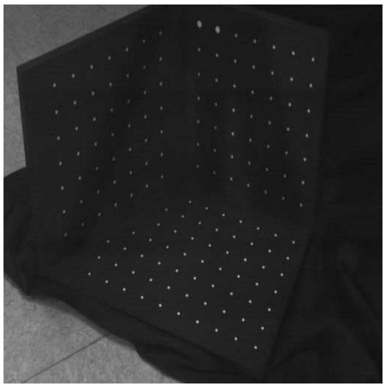


Figure 3: *The calibration grid, this grid is composed of retro-reflective targets.*

ΛP_1 than with ΛP_2.

$$\Lambda P_1 \leq \Lambda P_2 \quad , \quad \sigma_f^{\Lambda P_1} \leq \sigma_f^{\Lambda P_2} \ . \tag{11}$$

The maximum eigenvalue λ_{max} can also be determined from

$$\Lambda P q = \lambda q \ . \tag{12}$$

The square root $\sqrt{\lambda_{max}}$ is related to the maximum standard deviation. Another metric for comparing covariance matrices ΛP and ΛQ (an ideal matrix) can be achieved by using the sum of the squared logarithms of the eigenvalues (see Förstner [För95]):

$$d(\Lambda P, \Lambda Q) = \sqrt{\sum_{i=1}^{3} \ln^2 \lambda_i(\Lambda P, \Lambda Q)} \ . \tag{13}$$

However, this metric requires a criterion matrix and due to the methodology, we cannot establish an ideal matrix, because we do not know the answer in advance. In this way, the maximum eigenvalue of ΛP may be replaced by a less tight norm in order to avoid the rigorous determination of the maximum eigenvalue, e. g. the trace

$$tr(\Lambda P) = \sum_{i=1}^{3} \Lambda P_{ii} = \sum_{i=1}^{3} \lambda_i(\Lambda P) \ . \tag{14}$$

However, due to the fact that the maximum eigenvalue is in relation to the maximum standard deviation. Instead of the trace of ΛP we propose to use as a measure $\mu(P)$ the *maximum element in the diagonal of* ΛP which corresponds to the worst variance among the individual components

$$\mu(P) = \max_{j=1\ldots3} \Lambda P_{jj} \ . \tag{15}$$

This measure is easy to compute and gives a good uniformity of network precision in each coordinate axis of the error ellipsoid represented by ΛP.

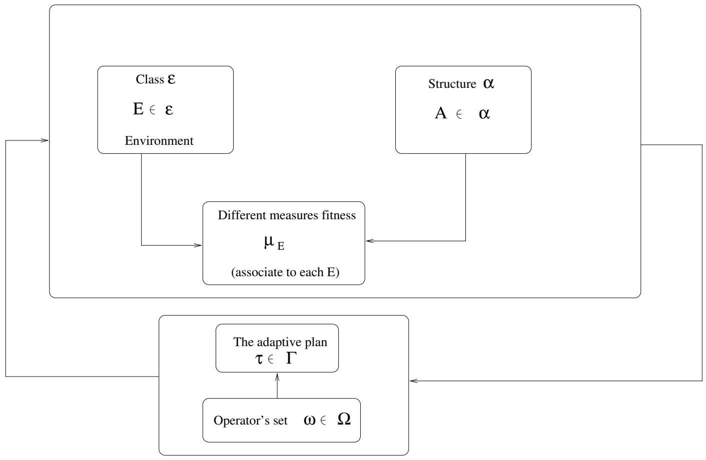


Figure 4: *The adaptive system.*

2.5 The Optimization Process

The problem of camera network design presents discontinuitie aspects mainly due to the unobservability of points which leads to a combinatorial optimization process which we have approached using a multicellular genetic algorithm. Genetic algorithms are probabilistic parallel search techniques based on the mechanism of natural selection and natural genetics [Gol89]. Since their development in the late 1960s [Hol92], genetic algorithms have been proven effective in searching large, non-linear, complex and poorly-understood search spaces, where expert knowledge is scarce or difficult to encode and where traditional optimization techniques fail. They are capable of solving non-linear problems. Because they do not require derivative information to direct their search, they are not limited to problems that are continuous. Instead of progressing from point to point, like other techniques, genetic algorithms search from a set of problem solutions to another. This feature allows them to escape local optimum, making their search more global in nature.

2.5.1 The Adaptive System

The camera network design can be achieved following genetic algorithm methodology. This methodology is composed of five major components, see Figure 4.

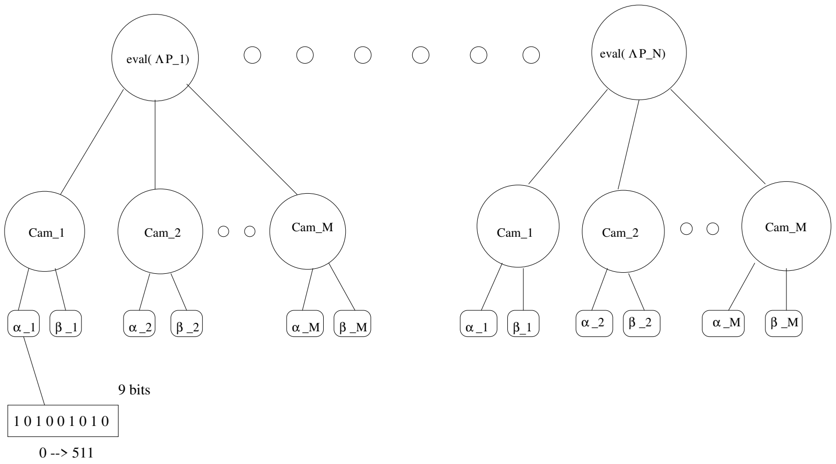


Figure 5: *The multicellular genetic algorithm is represented by a tree structure composed of a main node where the evaluation process is stored and several leaves corresponding to each camera. All cameras are codified in two parameters* (α, β) *which corresponds to the cells of an artificial being.*

1. The definition of a structure, $A \in \alpha$, which represents a tentative solution to the problem. It is represented here as a set of variables which are grouped into just one common structure, see Figure 5.

2. The environment, $E \in \varepsilon$, which limits the structure. It is represented here as the set of geometrical and optical constraints.

3. A measure μ_E of performance, i. e., the fitness of the structures for the environment. It is represented here as the value $\mu(P)$

4. The adaptive plan, $\Gamma \in \tau$, whereby the system's structure is modified to effect improvements. This is the genetic algorithm detailed later on.

5. The operator's set, $\Omega \in \omega$, which are used by the adaptive plan. This is represented here by the crossover and mutation operations.

These five main points are fundamental to establish a strategy for the camera network design problem. This section describes a multicellular genetic algorithm (MGA) used to solve the camera position and orientation problem in order to obtain the minimal 3D error. The idea is to create a computational model to optimize a set of convergent camera networks using an evolutionary technique. This model permits to implement the design as a problem of *optimization*. We propose a specialized implementation of the standard genetic algorithm approach, following the tree-based genetic programming representation [Koz92], [Kin94] which we assume here as known and fixed, see figure 5. This representation partitions the parameters into several distinct parts or "cells", each one corresponds to one camera variable. Each camera evolves only with cameras of the same class. This partition allows the emergence of subpopulations of good individuals as we will see in the experiments. As an example, all the variables of camera 1 of each tree structure are combined among them, and so on for the other cameras. As a matter of fact, some of these cells are then concatenated to form one long binary string. Binary representation is prefered here to float due to the fact that *two points close to each other in the representation space are not close to each other in the solution space.* This allows a random search in the whole space. In the other side, the probability to have a very large change in the value of the parameters, when you have changed a single bit, is rather small only in some high-order bits. In this way, a robust structure is obtained which performs well to the changes of the environment constraints. Our algorithm does not measure the fitness of each cell individually. Instead a global evaluation, before applying the geometrical and optical constraints, of a well designed criterion using derivatives is implemented. A tree structure represents a camera network configuration which is observing several parts of the object, groups of targets. This camera network can be divided into several subnetworks, each one is observing one group of targets. Then a genetic algorithm for each variable is executed, with just one general evaluation $\mu(P)$ among all the camera subnetworks. The fitness will be then the worst among these measures for a given object. Thus, we can see the algorithm as an idealized genetic algorithm IGA [Mit96, page 132] where each partition is sampled independently and the best schema, set of bits, in each partition tends to be selected.

To keep the search space reasonably small, we use the viewing sphere model, where the cameras move on a sphere looking inwards towards a central point. This model provides convergent configurations which give improved object measurement precision compared to other network configurations [Fra96]. In this way, cameras move in the space to achieve the design goals. The direct relationship between the spherical coordinates of the viewing sphere and the components of imaging geometry permits the heuristic search of the genetic algorithm to be exploited. Consequently, the spherical coordinate system (α, β, r) is well suited to the representation of the search space. All cameras are codified by the two parameters α and β, as r, the distance between the central point and the cameras, is considered as a constant here. If the environment constraints as field of view and resolution are taken into account, the variable r, will change. However, this do not change the final network configuration found by our system. Thus, the MGA is represented by a fixed tree structure composed of a main node where the evaluation process (the maximum value in the diagonal of ΛP) is stored and several leaves corresponding to each camera.

2.5.2 The Multicellular Genetic Algorithm

The multicellular genetic algorithm then proceeds as follows:

1. an initial random population of N convergent networks that satisfy the environment constraints is chosen and it is represented by (α_n, β_n) coded into a binary string representation.

2. next, we evaluate each network, and store the corresponding maximum value of the diagonal of ΛP_n for each tree structure. This corresponds to the fitness value which says how good the network is, compared with other solutions in the population $P(t)$.

3. then, we select a population of "good" networks by *tournament selection* [BT95]: two networks are selected from $P(t)$ and are compared selecting the best individual according to its fitness yielding the population $P(t+1)$.

4. from this population, we recombine the binary strings (α_n, β_n) for each camera using the following operations.

- Crossover, with a probability[3] Pc = 0.7, this operation was implemented using *one-cut-point*[4]. Let the two parents be:

$$\alpha_x = [\alpha_{x1}\ \alpha_{x2}\ \alpha_{x3}\ \alpha_{x4}\ \alpha_{x5}\ \alpha_{x6}\ \alpha_{x7}\ \alpha_{x8}\ \alpha_{x9}] \ ,$$

$$\alpha_y = [\alpha_{y1}\ \alpha_{y2}\ \alpha_{y3}\ \alpha_{y4}\ \alpha_{y5}\ \alpha_{y6}\ \alpha_{y7}\ \alpha_{y8}\ \alpha_{y9}] \ ,$$

  if they are crossed after the random kth position $= 4$, the resulting offsprings are:

$$\alpha'_x = [\alpha_{x1}\ \alpha_{x2}\ \alpha_{x3}\ \alpha_{x4}\ \alpha_{y5}\ \alpha_{y6}\ \alpha_{y7}\ \alpha_{y8}\ \alpha_{y9}] \ ,$$

$$\alpha'_y = [\alpha_{y1}\ \alpha_{y2}\ \alpha_{y3}\ \alpha_{y4}\ \alpha_{x5}\ \alpha_{x6}\ \alpha_{x7}\ \alpha_{x8}\ \alpha_{x9}] \ .$$

- Mutation; with a probability Pm = 0.005, this operation alters one or more genes. Assume that $(\alpha_{y5} = 1)$ gene of the chromosome α'_x is selected for a mutation. Since the gene is 1, it would be flipped into 0.

These operations yields a new population which we copy into $P(t)$.

5. Steps 2,3, and 4 are repeated until the optimization criterion stabilizes.

Finally, this algorithm minimizes the maximum value in the diagonal of ΛP

$$\min_{i=1 \ldots N} (\mu(P)) \ . \tag{16}$$

Thereby, the camera placement M_i relative to the world coordinate frame is optimized. Geometrically, each ΛP_i represents a hyper-ellipsoid, which changes its orientation and size as each sensor placement M_i does. Thus, an optimal placement solution is proposed, where the combined uncertainty of all points is minimal.

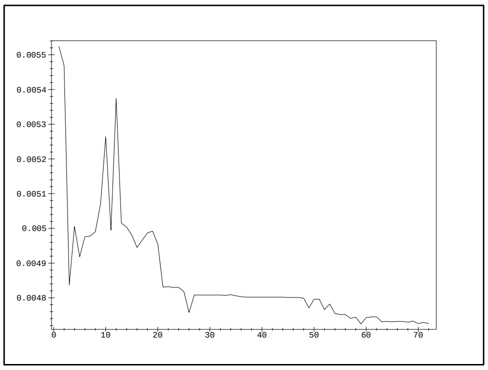
[3] For a discussion of the threshold values associated to Pc and Pm see [Mit96].

[4] due to the classification of the MGA this operation works like a *multiple-cut-point.*

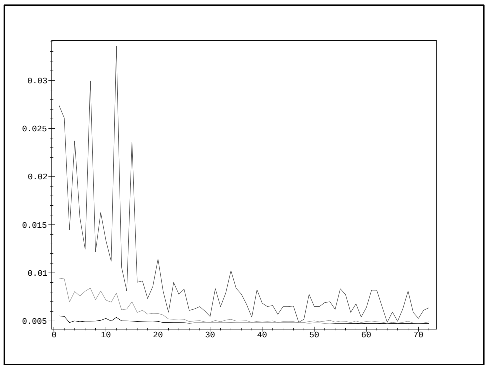
3 Experiments

We have run a series of experiments to test the validity of our approach. We present some results in Figure 11 which shows six cases of configurations designed by EPOCA in which several cameras are looking at several targets represented by their error ellipsoids aligned in one, two, or three planes. These configurations are product of our evolutionary system.

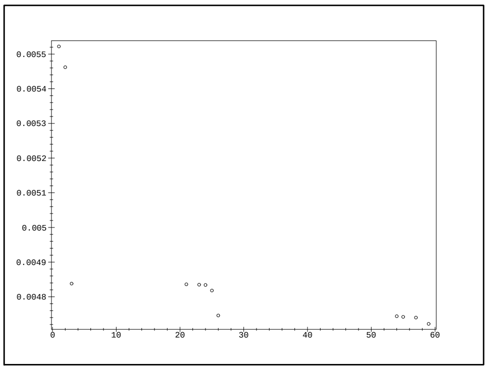
As an example the graphs of Figure 6 were produced with a population of 30 networks composed of four cameras which converged in no more than 150 generations. These graphs show the maximal, average and best fitness for one run of the algorithm. These results are similar to the four-camera network of Figure 11b. This experiment took a total of 72 generations to converge, discovering the best configuration with a fitness value $= 4.7 \times 10^{-3}$ in generation 59. It has required about 7 seconds of CPU time on an UltraSPARC 200 Mhz. Figure 6c shows the moment where a minimum appears along the run. Note also that evolution is terminated near generation 60.



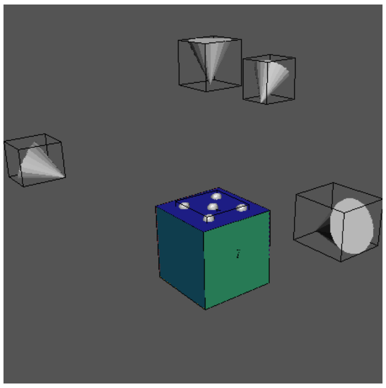
a) The plot of best fitness.



b) Maximal, average and best fitness.



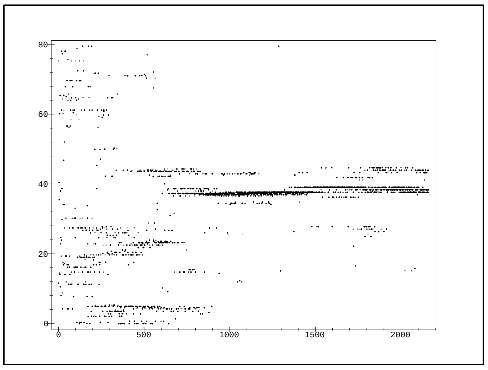
c) The last best configuration.



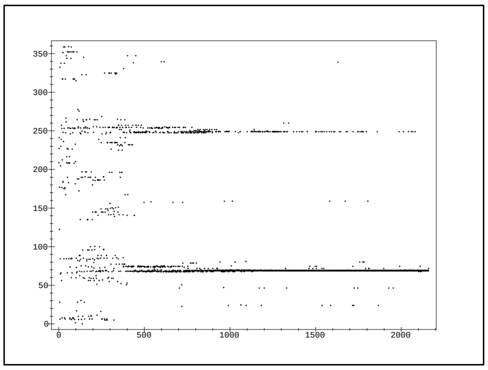
d) The best configuration.

Results of the evolution process for a single MGA run.

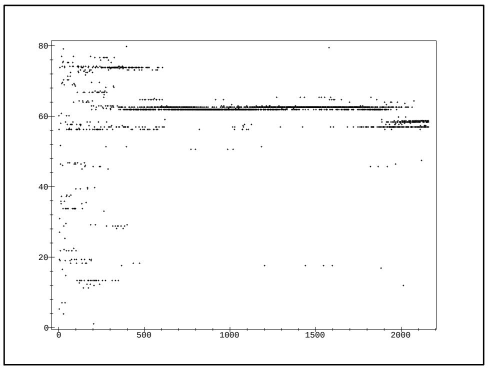
Figure 6: *The following graphs show the statistics and convergence time of a single MGA run composed of 4 convergent cameras.*



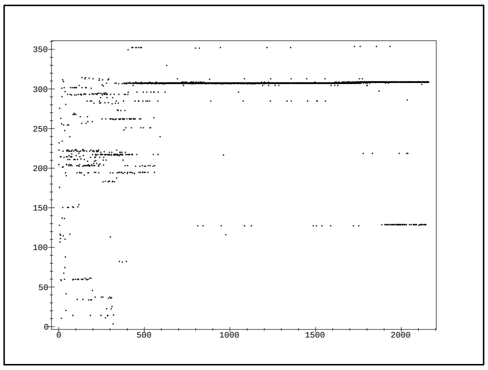
a) Angle α of camera 1.



b) Angle β of camera 1.



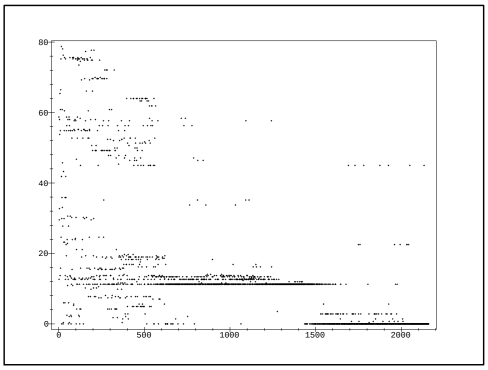
c) Angle α of camera 2.



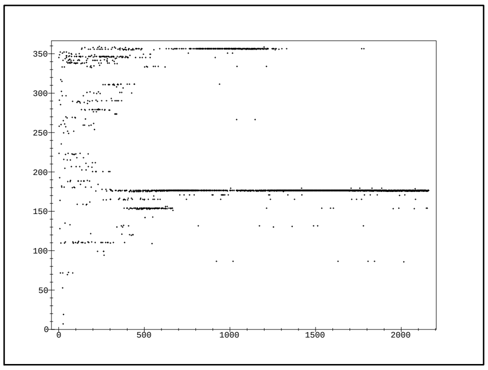
d) Angle β of camera 2.

Results of the evolution process applied to the camera variables.

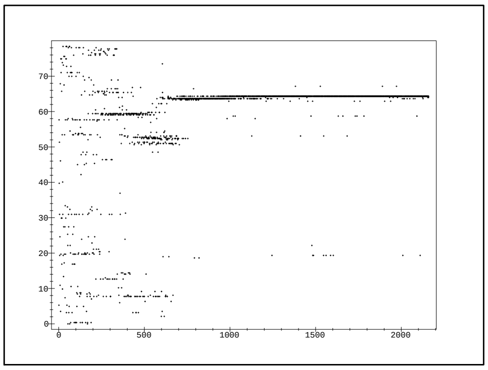
Figure 7: *The following graphs show the evolution of cameras 1 and 2 represented by their variable positions (angles α and β). Y axis is in degrees and X axis shows the number of the considered sample. It is important to notice that each graph is divided in 72 generations each one containing 30 samples, which represents 2160 samples.*



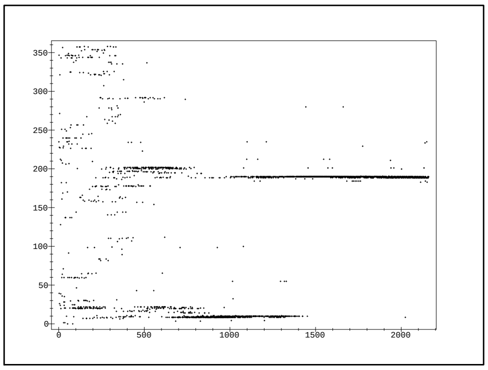
a) Angle α of camera 3.



b) Angle β of camera 3.



c) Angle α of camera 4.

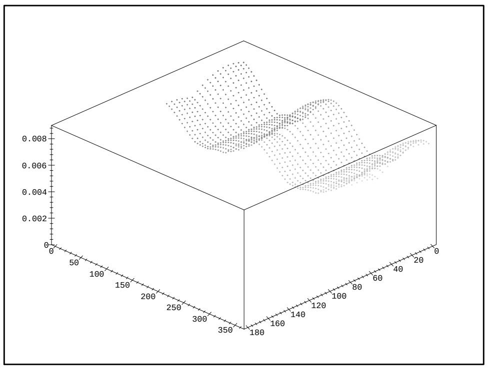


d) Angle β of camera 4.

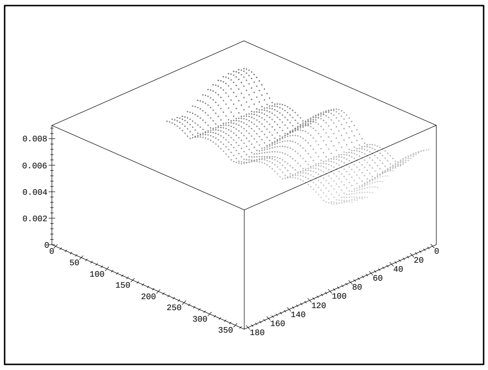
Results of the evolution process applied to the camera variables.

Figure 8: *The following graphs show the evolution of cameras 3 and 4 represented by their variable positions (angles α and β). Y axis is in degrees and X axis shows the number of the considered sample. It is important to notice that each graph is divided in 72 generations each one containing 30 samples, which represents 2160 samples.*

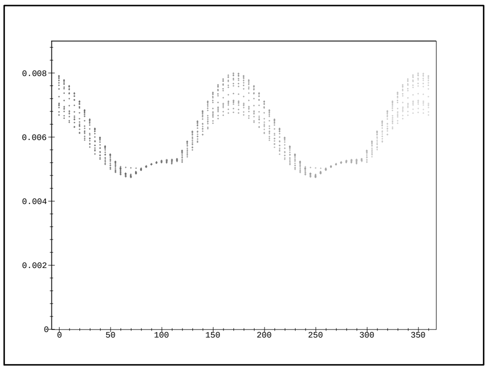
Landscape graphs of camera 1 and 2.



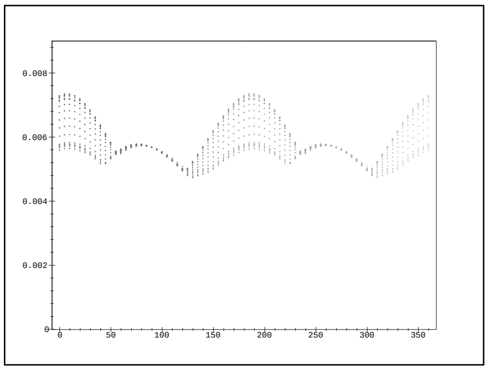
a) Surface graph of camera 1.



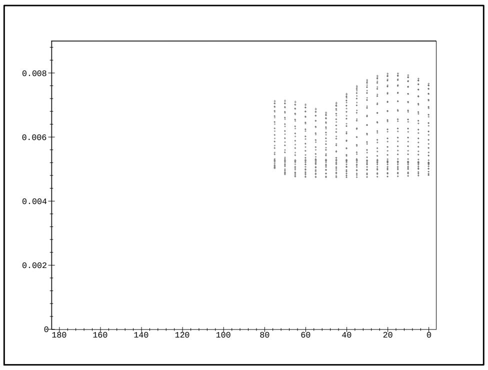
b) Surface graph of camera 2.



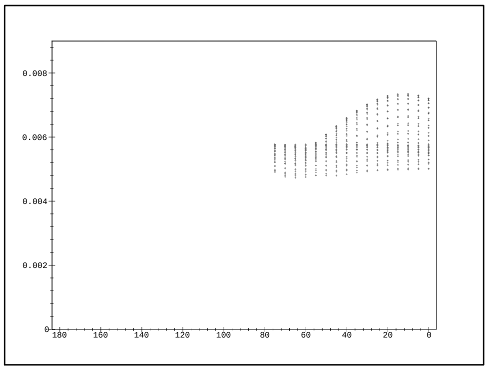
c) Frontal view of graph a.



d) Frontal view of graph b.



e) Lateral view of graph a.

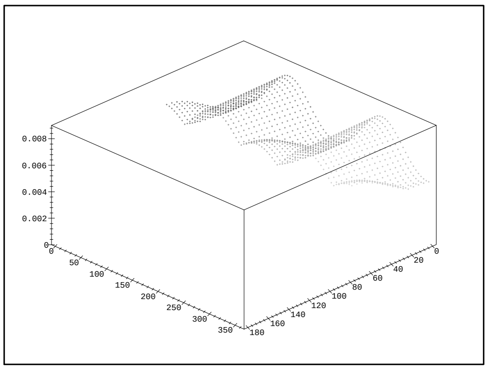


f) Lateral view of graph b.

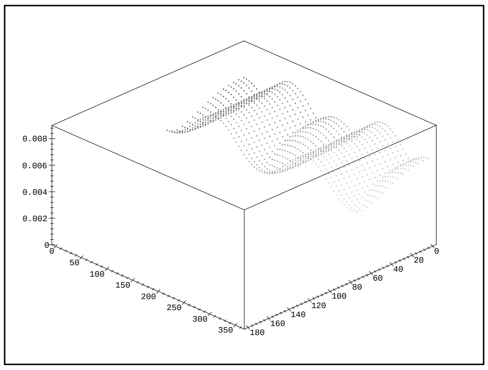
Final landscape's results of the four-camera network.

Figure 9: *The graphs on the left correspond to the landscape of camera 1 while the graphs on the right correspond to the landscape of camera 2.*

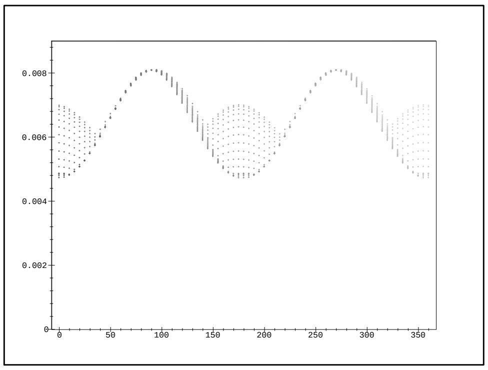
Landscape graphs of camera 3 and 4.



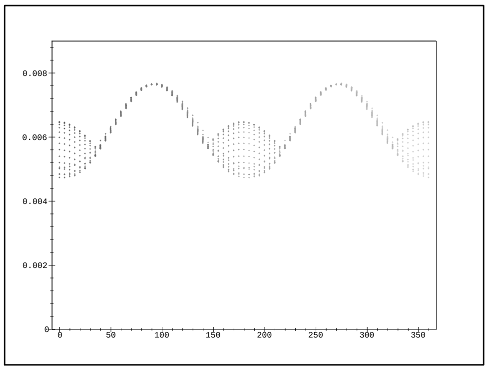
g) Surface graph of camera 3.



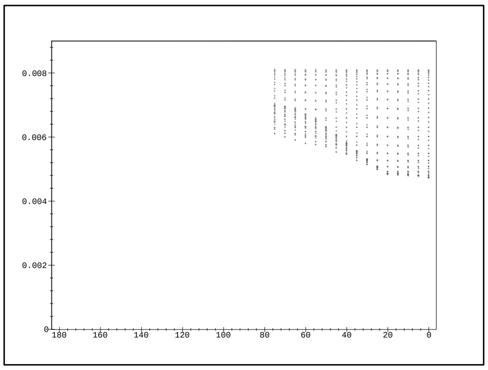
h) Surface graph of camera 4.



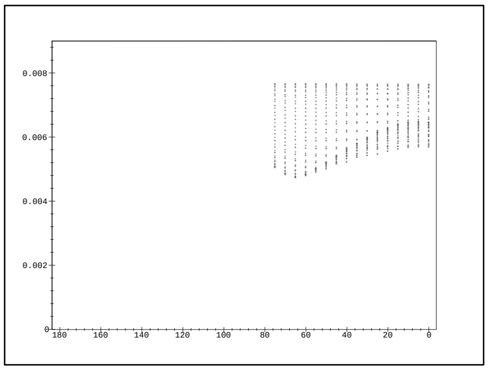
i) Frontal view of graph g.



j) Frontal view of graph h.



k) Lateral view of graph g.



l) Lateral view of graph h.

Final landscape's results of the four-camera network.

Figure 10: *The graphs on the left correspond to the landscape of camera 3 while the graphs on the right correspond to the landscape of camera 4.*

Camera	Angle α	Angle β
1	39.02	68.90
2	62.92	308.67
3	0.00	176.48
4	64.33	184.21

Table 2: Final results of the four-camera network.

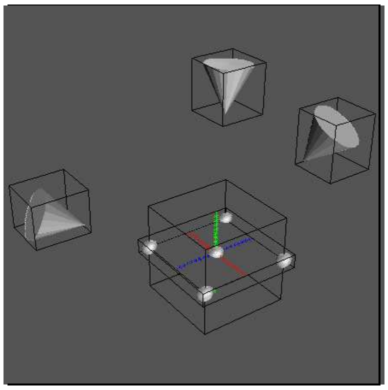
Table 2 presents the final results of the four-camera network configuration. Notice the behaviour of the camera population as illustrated in Figures 7 and 8. We see that after a random distribution, emergent coherent configurations appear. For instance, the α angle of camera 1, see Figure 7a, converges towards two dominating values below and above 40 degrees, while the β angle in the same camera is about 69 degrees. This last has a second minimum about 250 degrees, see Figure 9c, 180 degrees after the first minimum. Other dominating configurations appear too, but disappear before half of the total iteration. Similar observations can be made on all other parameters.

If we analyse how $\mu(P)$ evolves with these parameters, we will observe the local minima of the dominating configurations. Figure 9 displays the graph of $\mu(P)$, after convergence, only with respect of two parameters and for camera 1 and 2. It clearly shows the two minima for the β parameter of camera 1, while the landscape is quite flat for the α parameter. Similar observations can be derived in the graphs of Figure 9. Cameras 2 and 4 have also two local minima, in contrast to camera 3 which presents just one minimum. The α angle is about 0 degrees (i. e., camera 3 is orthogonal to the surface plane), indeed this is a singular configuration for which the parameter β of camera 3 can have any angle.

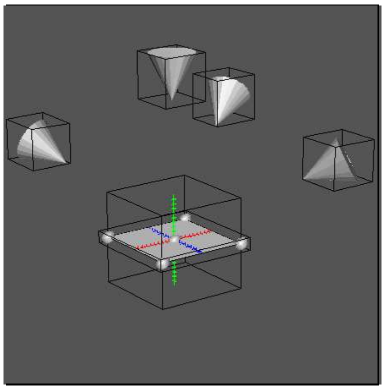
In general, the population size must be large enough to let the MGA find the good solutions. A small population (less than 10 networks) discourages premature convergence to suboptimal solutions[5] while a large population requires more evaluations per generation to achieve the same fitness values. The problem of convergence is usually related to the length of the string [MFH93].

[5]This is mainly due to the need of niches which are landscapes from which the global or optimal solution is chosen. Consequently, if these zones are less rich, the opportunities to obtain a better design are reduced.

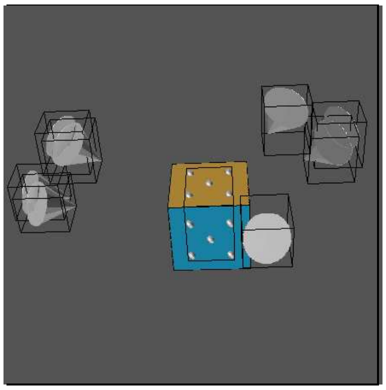
EPOCA network designs.



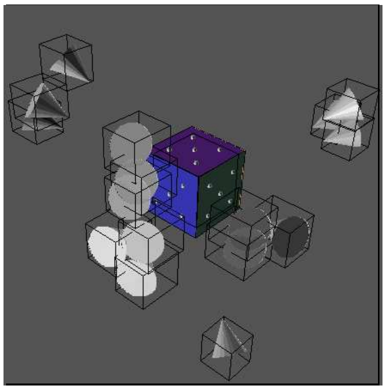
a) 3 cameras observing a plane.



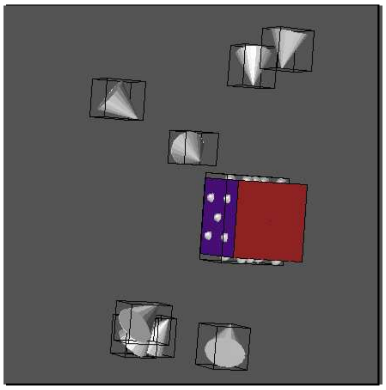
b) 4 cameras over a plane.



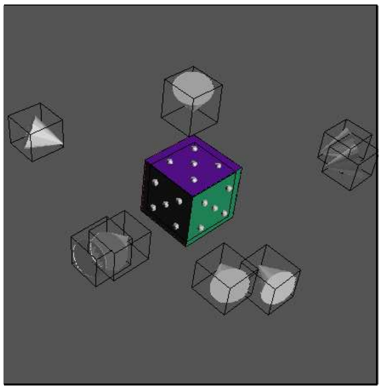
c) 8 cameras over 2 adjoining planes.



d) 16 cameras over 3 adjoining faces.



e) 8 cameras observing 3 opposite faces.



f) 8 cameras over 3 adjoining faces.

Examples of configurations designed by EPOCA.

Figure 11: *Between the several designs proposed by EPOCA we have identified b) like one used by Fraser [Fra82] which is not atypical of an imaging geometry.*

In our implementation strings of 300 or more bits have been optimized thanks to the tree classification.

Other examples for 2 and 3 adjoining planes are presented in Figure 11. These examples show how the cameras are positioned at very precise zones of the space for which the overall uncertainty is minimal. Note that in the cases of 3 planes, the cameras are not observing all targets at the same time.

Figure 11b illustrates a solution with four cameras looking at a planar surface. This solution is not the standard one used by the experts: photogrammetrists usually put the four cameras at four corners of a cube whose center contains the targets to be measured. In fact our configuration has already been discussed by Fraser [Fra82]; he noticed that this configuration is not atypical. Our experiments confirm Fraser's statement, we even claim that this configuration is slightly better than the classical one.

4 Discussion and Perspectives

In this paper we have set the problem of camera placement with the goal to achieve accurate 3D measurements in terms of an optimization design. The problem has been divided into two main parts. The first, was dedicated to an analytical error model from which a criterion was derived. The criterion we have chosen is the *maximum element in the diagonal*, extracted from the covariance matrix which is based in the error propagation phenomenon. This criterion was chosen in order to speed up the MGA process.

The second part, was dedicated to a global optimization method which has minimized the criterion. Due to the unobservability of points, caused by the different constraints, the problem presents discontinuities which leads to combinatorial aspects in the optimization process. These constraints are naturally incorporated to the genetic algorithm methodology. Our idea was to create an artificial system modelling the uncertainty of the reconstruction process. We can see the system like an information processing system, rich in autonomy and creativity, due to the evolutionary technique we have used. In this way, the system is able to evolve its own structure autonomously. This is an adaptive complex system that is similar to living systems. The system is a dynamic process where the global state *emerges* by coevolution through the behavior and interaction of the components "cells" which are related by coadaptation. It is necessary to say that the behavior of these cells is influenced by the global state. An extrinsic parallelism, together with the intrinsic parallelism, is executed. A phenomenon of *niche* is present in each camera representation. These niches evolve with respect to the others mainly due to the rules of the system. Our experience confirms results obtained by other researchers [BATM95]; genetic algorithms are able to solve efficiently such kind of optimization problems with high combinatorial aspects.

Our EPOCA system succesfully produces two and three camera networks designs similar to those used by photogrammetrists. In the case of four cameras a non-standard design was proposed which should give slightly better results compared to the more classical networks. Moreover, the system can design networks for several adjoining planes. All the configurations are good in terms of the camera distribution and ray inclination.

This work has to be considered as an additional step towards automated camera network design. The strength of our approach is its generality. This work could easily be extended by considering alternative optimization process working on a non uniform criterion. However, several simplifying assumptions were made here that have to be explored in the future. First, we did not have complex occlusions, nor we had complex constraints on camera position. This kind of additional constraints will increase the combinatorial complexity aspect of the optimization and the open questions that arise are: Will such system converge? How long will it take? A second point is that we assume the camera external and internal parameters are perfectly known. In fact the user has good estimates for them, but if high accuracy is needed, the bundle adjustment would have to refine their estimates together with the 3D measurement estimation. This complicates the criterion computation at each evaluation step in the optimization. If this might not be a big theoretical issue, it is a practical one as the computation time could be multiplied by a huge scale factor. These two aspects are both related to the computational complexity of the optimization; to explore them is our goal for future research.

Acknowledgments Gustavo Olague is grateful for research founding of CONACYT México (grant 66291). Figures 1, 6d and 11 were generated with software written at the Geometry Center, University of Minnesota. We are also grateful to Dr. Scott Mason for his helpful comments.

References

[AAT] S. Abrams, P.K. Allen, and K.A. Tarabanis. Dynamic sensor planning. unpublished.

[BATM95] P. Bessière, J.M. Ahuactzin, E. Talbi, and E. Mazer. The ariadnés clew algorithm: Global planning with local methods. In Latombe Goldberg, Halperin and Wilson, editors, *Algorithm Foundations of Robotics*, Boston, MA, 1995. A. K. Peters.

[Bey92] H.A. Beyer. *Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System.* PhD thesis, ETH-Zurich, 1992.

[BT95] T. Blickle and L. Thiele. A mathematical analysis of tournament selection. In *Proceedings of the Sixth International Conference on GA.*, pages 9–16. Univ. of Pittsburgh, July 1995.

[CK88] C.K. Cowan and P.D. Kovesi. Automatic sensor placement from vision task requirements. IEEE *Transactions on Pattern Analysis and Machine Intelligence*, 10(3):407–416, May 1988.

[För95] W. Förstner. A metric for comparing symmetric positive definite matrices. unpublished, August 1995.

[Fau93] O. Faugeras. *Three-Dimensional Computer Vision - A Geometric Viewpoint.* Artificial intelligence. The MIT Press, Cambridge, MA, USA, Cambridge, MA, 1993.

[FC90] D. Fritsch and F. Crosilla. First order design strategies for industrial photogrammetry. In *Close-Range Photogrammetry Meets Machine Vision.*, volume 1395, pages 432–438. SPIE - Society of Photo-Optical Instrumentation Engineers, 1990.

[Fra82] C.S. Fraser. Optimization of precision in close-range photogrammetry. *Photogrammetric Engineering and Remote Sensing*, 48(4):561–570, April 1982.

[Fra92] C.S. Fraser. Photogrammetric measurement to one part in a million. *Photogrammetric Engineering and Remote Sensing*, 58(3):305–310, March 1992.

[Fra96] C.S. Fraser. Network design. In K.B. Atkinson, editor, *Close Range Photogrammetry and Machine Vision*, chapter 9, pages 256–281. Whittles Publishing, Roseleigh House, Latheronwheel, Caithness, KW5 6DW, Scotland, UK, 1996.

[Gol89] D.E. Goldberg. *Genetic Algorithms in Search Optimization and Machine Learning.* Addison-Wesley, 1989.

[Gru85] A.W. Gruen. Adaptative least squares correlation: a powerful image matching technique. *S. Afr. Journal of Photogrammetry, Remote Sensing and Cartography*, 14(3):175–187, 1985.

[Hol92] J.H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analisis with Applications to Biology, Control and Artificial Intelligence.* The MIT Press, Cambridge, MA, USA, 1992.

[Kin94] K.E. Kinnear. *Advances in Genetic Programming.* The MIT Press, Cambridge, MA, USA, 1994.

[Koz92] J.R. Koza. *Genetic Programming, On the Programming of Computers by Means of Natural Selection.* The MIT Press, Cambridge, MA, USA, 1992.

[Mas97] S. Mason. Heuristic reasoning strategy for automated sensor placement. *Photogrammetric Engineering and Remote Sensing*, 63(9):1093–1102, September 1997.

[MFH93] M. Mitchell, S. Forrest, and J.H. Holland. The royal road for genetic algorithms: Fitness landscapes and ga performance. In F.J. Valera and P. Bourgine, editors, *Proceedings of the First European*

Conference on Artificial Life, pages 245–254. MIT Press/Bradford Books, 1993.

[MG95] S.O. Mason and A. Gruen. Automatic sensor placement for accurate dimensional inspection. *Computer Vision and Image Understanding*, 3(61):454–467, 1995.

[Mit96] M. Mitchell. *An Introduction to Genetic Algorithms.* The MIT Press, Cambridge, MA, USA, second printing edition, 1996.

[Moh93] R. Mohr. Projective geometry and computer vision. In C.H. Chen, L.F.Pau, and S.P. Wang, editors, *Handbook of Pattern Recognition and Computer Vision*. World Scientific Pub., 1993.

[PFTV88] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C.* Cambridge University Press, 1988.

[SNSS92] S. Sakane, R. Niepold, T. Sato, and T. Shirai. Illumination setup planning for a hand-eye system based on an environmental model. *Advanced Robotics*, 6(4):461–482, 1992.

[SS91] S. Sakane and T. Sato. Automatic planning of light source and camera placement for an active photometric stereo system. In *Proceedings of* IEEE *International Conference on Robotics and Automation, Sacramento, California, USA*, pages 1080–1087. IEEE Robotics and Automation Society, April 1991.

[TG93] G.H. Tarbox and S.N. Gottschlich. Planning for complete sensor coverage in inspection. Technical Report TR-CAT-93-4, NYS Center for Advanced Technology in Automation and Robotics., 1993.

[TTA91] K. Tarabanis, R.Y. Tsai, and P.K. Allen. Automated sensor planning for robotic vision tasks. In *Proceedings of* IEEE *International Conference on Robotics and Automation, Sacramento, California, USA*, pages 76–82, April 1991.

[YHS95] S. Yi, R.M. Haralick, and L.G. Shapiro. Optimal sensor and light source positioning for machine vision. *Computer Vision and Image Understanding*, 61(1):122–137, 1995.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399