Soft Comput (2008) 12:751–764
DOI 10.1007/s00500-007-0234-3

ORIGINAL PAPER

Fuzzy-genetic optimization of the parameters of a low cost system for the optical measurement of several dimensions of vehicles

J. Otero · L. Sánchez · J. Alcalá-Fdez

Published online: 11 September 2007

Abstract When designing optical measurement systems, it is common to use cameras, lenses and frame grabbers specially designed for metrology applications. These devices are expensive, therefore optical metrology is not the technology of choice in low cost applications. On the contrary to this, surveillance video cameras and home oriented frame grabbers are cheap, but imprecise. Their use introduces inaccuracies in the measurements, that sometimes can be compensated by software. Following this last approach, in this paper it is proposed to use fuzzy techniques to exploit the tolerance for imprecision of a practical metrology application (to automate the measurement of vehicle dimensions in Technical Inspection of Vehicles in Spain, the equivalent of the Ministry Of Transport Test or MOT Test in UK) and to find an economic solution. It will be shown that a genetic algorithm (GA), guided by a fuzzy characterization of the sources of error, can optimize the placement of the video cameras in a station so that these mentioned sensors can be used to take measurements within the required tolerance.

Supported by the Spanish Ministry of Science and Technology under Projects TIN-2005-08386-C05-05 and TIN-2005-08386-C05-01.

J. Otero · L. Sánchez (✉)
Department of Computer Science, University of Oviedo,
33204 Gijón, Spain
e-mail: luciano@uniovi.es

J. Otero
e-mail: jotero@uniovi.es

J. Alcalá-Fdez
Department of Computer Science, University of Jaén,
23071 Jaén, Spain
e-mail: jalcala@decsai.ugr.es

Keywords Fuzzy uncertainty · Genetic algorithms · Random sets · Stereoscopic vision · Metrology

1 Introduction

Measurement of dimensions (width, length, height, wheelbase, ramps and other accessories) is one of the tasks to be performed by companies authorized to carry out Technical Inspection of Vehicles (TIV) in Spain (the equivalent of the Ministry Of Transport Test or MOT Test in UK). The standard procedure requires two human operators which makes it a costly process. Unfortunately, although its automatization is not a conceptually complex problem, neither the cost of the hardware that should be acquired, nor the alterations in the disposition of the TIV station that would allow the automatic measurement of dimensions, are affordable.

Since hardware cost is the main constraint, a partial automatization of the measuring process, based on cheap general purpose cameras, was considered. A prototype of such system has been implemented in the TIV of Pruvia (Asturias, Spain) and is shown in Fig. 1. To reduce costs, the system consists of four cameras of the type usually found in surveillance applications. Furthermore, images are taken sequentially, because only one capture board, with a single channel, was used.

In Fig. 2 the system is depicted. The computer is a standard PC with a Pentium IV processor at 2.4 GHz, 512 GB RAM, running Red Hat Linux 9 (Shrike). The input of a single channel capture board (an AVerTV from AVerMedia with a maximum resolution of 768 × 576 pixels) is connected to the output of a multiplexer. This multiplexer is controlled by the parallel port of the computer after a voltage level conversion. In this way, the software can sequentially select one of the cameras (each one a Fujitsu TCZ-260E with Computar 8 mm

Fig. 1 Partial View of the Vehicle Measurement System

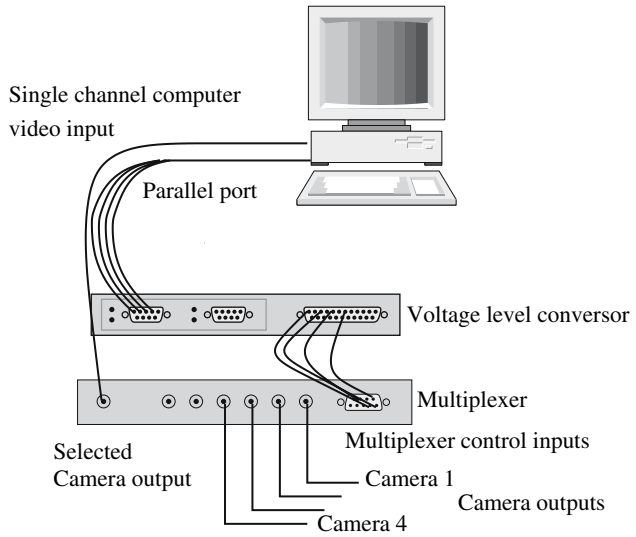

Fig. 2 Schema of the acquisition system. A single channel capture board is used. A multiplexer allows to select one camera at a time. The images are not taken at the same time

lens) at a time. With this schema, the whole set of four images can be obtained sequentially. The difference in acquisition time makes it possible that some objects in the scene and even the illumination could change. In fact the acquisition delay between two consecutive images is in the range of a few seconds due to the cable length and to the analogue nature of the cameras (see Fig. 3 bottom images, where it is shown how the pose of the technician that is working with the vehicle has changed).

The cameras were not mounted in pairs as usual in stereo applications (Brown et al. 2003), in fact they are some meters apart but they are physically positioned so that all of the points that define each measurement are in the field of view of at least two of the cameras. This arrangement still allows us to use stereoscopic vision techniques to find the spatial coordinates of the points involved, but prevents the application of classical matching algorithms (Bhat and Nayar 1998; Faugeras et al. 1993; Hoff and Ahuja 1989; Scharstein and Szeliski 1998; Tomasi and Manduchi 1998) to pairs of images. There are recent approaches (Santamaría et al. 2007) that use metaheuristics and genetic algorithms to find the best correspondences. In our case the images that form a stereo pair of are very different between them (the cameras are spaced about 10 meters apart, see Fig. 11) and the transformation of the view taken by a camera can hardly be matched with an image taken by another one. To our knowledge, no standard matching technique (see Brown et al. 2003 for an exhaustive compilation) can be applied with such extreme configuration, regardless of a previous rectification of the image (see top images in Fig. 3). Even more, due to the low-cost hardware used and the diversity of available models in the market, the acquired images do not allow to recognize automatically some accessories of the vehicles. For example, in the bottom images in Fig. 3 the points that define the maximum width of the vehicle (located at the driving mirrors) are shown in both images of a stereo pair (marked with + and white arrows). Only a skilled trained technician can find the same points in both images, as can be seen in Fig. 4, where the driving mirror on the driver side is shown enlarged.

However, this is not an important problem, since a human operator is needed to select the bounds of the vehicle that must be measured. Even more, it is mandatory the presence of the operator because legal issues, in order to certificate the measurements, including the verification of the points that define the measurement. With little extra effort, the same operator can select these points twice (once for each view of the point) obtaining a stereo pair effectively and complete both tasks (measurement and certification) at the same time and more efficiently.

The main problem with this semi-automatic procedure arises from the fact that the accuracy of the measurement is very sensitive to errors in the selection of the points on the computer screen. Furthermore, we have observed that the average error is strongly correlated with the relative situation of each camera with the other cameras, and the relative situation of all the cameras with the vehicle position. This problem is common to many stereo procedures (Duda and Hart 1973), and has been previously studied from a theoretical point of view. The relationship between image resolution and the precision in depth measurement was carried out in McVey and Lee (1982) and Verri and Torre (1986), where the effect of camera positioning and relative orientation of image planes is studied separately. An in-depth analysis of the error in stereo setup is performed in Blostein and Huang (1987), including the analytical derivation of probability distributions of the errors in the three directions of the tridimensional space. Since an ideal model of the camera was employed (no distortions or electronic noise were considered), the results are an upper bound of any stereo measurement setup. More recently, in Mason and Grünt (1995), expert knowledge was

Fig. 3 Whole set of vehicle images used by the system. Two consecutive images are taken as a stereo pair. Each stereo pair shares an image with the previous and the next one: the four images shown form three stereo pairs. As can be seen, using standard matching techniques in order to find corresponding points when measuring the maximum width of the vehicle its impossible. In the bottom images the points that define the distance between the driving mirrors are marked. In Fig. 4 a detail of the driver side driving mirror on the bottom right image (the area surrounded by a *rectangle*) is shown

Fig. 4 Detail of the portion surrounded by a rectangle in the bottom right image in Fig. 3. As can be seen it is impossible to automatically detect the driving mirror, only a well trained human eye can do that and there is a inherent lack of precision in the measurement process

incorporated to obtain the relationship between the camera position and the measurement precision. Furthermore, the most valuable contribution related to our work is (Olague and Mohr 2002), in where a multicellular genetic algorithm is used to discover the position of a multi-camera measurement setup, in a problem that bears some relationships with ours.

Unfortunately, these works were not directly applicable to our problem:

– All of these studies are based on a model of the camera as a simple pin-hole. We need a more complex model, where the imperfection of the lenses is taken into account, because we use general purpose cameras, where the distortion should not be neglected. In other words, previous works are restricted to the use of high quality sensors.
– In these papers, the position of the cameras was left unconstrained. In our case (see Figs. 1 and 3, where the actual TIV station is shown) it is clear that cameras can not be placed anywhere: their location must be compatible with other activities at the plant, and walls, doors and machinery also restrict the feasible positions of the sensors.

Given the restrictions of the system being designed, we begin this paper by studying the sources of error related to the use of cheap elements (see Sect. 2, "Sources of error in the measurement"). In the following sections, the major contributions of this paper are explained. In Sect. 3, "Fuzzy characterization of the error in the measurement", we justify the use of a fuzzy characterization of all sources of error, including the human error when marking points (that turned out to be the most important source of imprecision) and build a quality index, which depends on the position of the cameras. Fuzzy sets were regarded as coverage functions of random sets, and that decision allowed us to introduce concepts of fuzzy statistics in our quality index. In Sect. 4, "Camera positioning optimization", a genetic algorithm was used to find the most favorable position of the cameras in terms of an upper bound of the quality index defined in Sect. 3. In the same section, the measured accuracy is also compared to that reached with a manual positioning of the cameras. Lastly, in Sect. 5, conclusions about this study are drawn.

2 Sources of error in the measurement

According to our experience, and to other works found in the literature, we have identified four sources of error in the optical measurement procedure: incorrect camera calibration, distortion of the lenses, quantization error, and human errors due to stereo pair determination. We are aware that two of these (distortion of the lenses, and quantization error, which will be discussed in Sect. 2.3) can be immediately reduced with a higher investment in hardware. Moreover, the camera calibration could also be partly automated, and the human errors reduced, if the number of cameras is doubled. In this case the determination of the stereo pairs could then be done by means of the matching algorithms mentioned in the introduction [see Brown et al. (2003) for details] using a standard stereo configuration where the cameras are placed in pairs. Since our objective was to develop a low cost application, these approaches were not followed given that they imply more investment in hardware (doubling the number of cameras and using a multiple channel frame grabber with greater resolution). Even more, as we will explain later in Sect. 3, if the cameras are closer, the uncertainty volume that contains a tridimensional point obtained from the bidimensional projection in the images, is larger (as can be seen in Fig. 8).

2.1 Camera calibration

In Brown et al. (2003) camera calibration is defined as *"Calibration is the process of determining camera system external geometry* (*the relative positions and orientations of each camera*) *and internal geometry* (*focal lengths, optical centers, and lens distortions*)*"* , in other words, it is the process of searching for the camera parameters that let us know with accuracy the points of the image where every visible point of the scene is projected (see Fig. 5). There are several approaches to realize effective camera calibration, we refer the reader to Faugeras and Luong (2001), Hartley and Zisserman (2000) and Olague and Dunn (2006).

Some of these parameters (like the focal length) are given by the manufacturer of the camera. Other parameters (like lens distortion) are not fully known and other parameters (like focus or zoom if available) can be changed by the user and interact with the former. In practice a model of the camera must be made from a set of points whose three-dimensional coordinates and image coordinates are known. Once a model is available, reverse calculation can be applied to know to which point in the scene space a given point in the image plane corresponds.

It is usual to establish the relation between the space coordinates and the image coordinates by eleven parameters, known as DLT (Direct Linear Transform) parameters

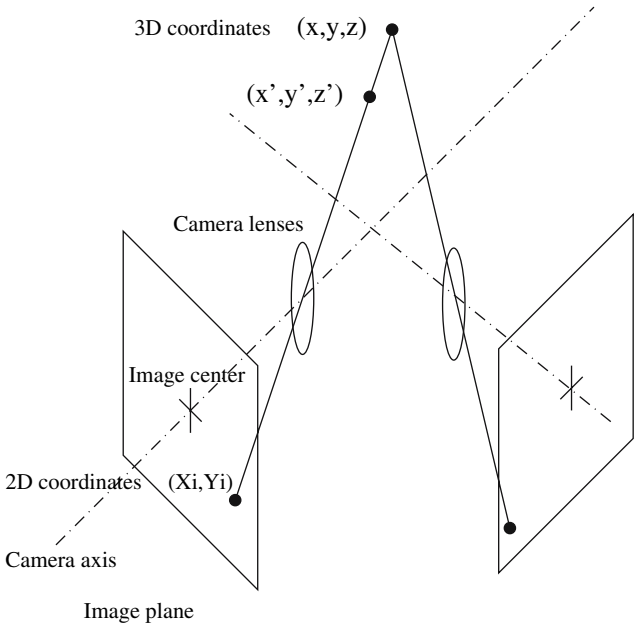

Fig. 5 Transformation from a point in the scene space O to a point in the image plane i. Each camera defines a line in the scene space, then at least two cameras are needed to calculate the point coordinates

by the creators of this method (Abdel-Aziz and Karara 1971; Marzan and Karara 1975), known as "resection" in other works (Olague and Dunn 2006). This model has been progressively improved by other researchers, see Hatze (1988) or Hinrichs and McLean (1995) for details. We are aware that other calibration methods exist (Faugeras and Luong 2001), we choose this because it is well known in Computer Vision community and serves as a testbed for the purposes of this paper. In this model, the relationship between the tridimensional coordinates and the coordinates in the image plane are given by the set of equations that follows:

$$\begin{aligned} x + \delta x &= \frac{L_1 X + L_2 Y + L_3 Z + L_4}{L_9 X + L_{10} Y + L_{11} Z + 1} \\ y + \delta y &= \frac{L_5 X + L_6 Y + L_7 Z + L_8}{L_9 X + L_{10} Y + L_{11} Z + 1} \end{aligned} \tag{1}$$

where X, Y, Z are the tridimensional coordinates of a point in the space and x, y are the bidimensional coordinates in the image coordinate system. The quantities δx and δy are the nonlinear errors in x and y directions. The parameters L_i, for i in $1 \ldots 11$ are the eleven DLT parameters that must be estimated from a set of points from which the tridimensional coordinates and the corresponding bidimensional coordinates are known.

The target of the calibration process is to obtain the eleven parameters for every camera. This process is applied to a set of tridimensional coordinates and the corresponding bidimensional coordinates in the camera being calibrated. For each pair of three and bidimensional coordinates, Eq. (2)

holds:

$$\begin{bmatrix} X & Y & Z & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & X & Y & Z & 1 \\ -xX & -xY & -xZ & & & & & \\ -yX & -yY & -yZ & & & & & \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \dots \\ L_{10} \\ L_{11} \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

Since there are 11 unknowns, at least 11 equations are needed. In practice, the minimum is 12, because from each pair of coordinates, two equations are needed. This leads to an overdetermined system that can be solved using least squares. The minimum number of points with known tridimensional and bidimensional coordinates needed to perform the calibration of a single camera is six (of course, one randomly chosen equation can be discarded, because of this, some authors claim that only 5.5 points are needed Olague and Dunn 2006). In practice, more points are needed to avoid the influence of individual errors (Chen et al. 1994). For instance, a number of 12–20 points is suggested in Shapiro (1978).

The DLT model, as shown in Eq. (2) allows us to simulate the process of image acquisition. But, as the model can be inverted, it also permits to simulate the process of measurement. Equation (3) serves us to obtain the tridimensional coordinates from the bidimensional data. Since data from a single camera provides us with three unknowns and two equations, at least two cameras are needed. Each camera is associated to the following equations:

$$\begin{bmatrix} L_1 - xL_9 & L_2 - xL_{10} & L_3 - xL_{11} \\ L_5 - xL_9 & L_6 - xL_{10} & L_7 - xL_{11} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} L_4 - x \\ L_8 - y \end{bmatrix} \quad (3)$$

and therefore an overdetermined system is obtained. It is easy to follow that from the two bidimensional coordinates of the projections of the same tridimensional point in two cameras, the tridimensional coordinates of the point can be found, as shown in Fig. 5. Although, from a theoretical point of view, there are no errors in this approach, some facts that lead to a lack of accuracy in the measurements must be taken into account when it is applied in practice. The set of parameters $L_1 \dots L_{11}$ that characterizes a camera is not invariant in time: it may be affected by changes in temperature, and the position of the cameras can be slightly modified by vibrations produced by the use of heavy machinery in the plant, which can not be regarded as completely rigid for the ranges of values involved. This means that the cameras should be recalibrated from time to time.

2.2 Distortion of the lenses

The lenses used in the system are not specially designed and manufactured to be used in metrology systems, so their defects have some influence on the accuracy of the measurements when the calibration points do not cover the whole plant due to the non-linearity of the true underlying model. There are several models to keep track of the effects of distortion and decentering of the lenses. They depend on more than eleven parameters (Hatze 1988; Hinrichs and McLean 1995) estimated from the views (from several cameras) of points whose tridimensional coordinates are known. It is common practice to use even more points than those suggested in Shapiro (1978). The minimum needed to obtain an acceptable error is 30 points, as pointed out in Hatze (1988).

In our case we chose to overcome the asymmetries introduced by the lenses by spreading the calibration points over the plant. A white cube made of wire (which can be seen in the right part of Fig. 1) was repeatedly placed across the plant and measured, to obtain a grid of calibration points. We could not find a set of DLT parameters uniformly valid for the problem, neither did we find that any of the mentioned models improved the quality of the measurement in a significant quantity. Instead, we divided the plant into four spatial volumes, each one of which was associated to a set of DLT parameters following an approach similar to the one in Kwon and Lindley (2000). When the prototype was put in production, each measurement was calculated four times, one for each set of parameters, and the spatial coordinates obtained by averaging the results, with a weight inversely proportional to the distances between the point being measured and the centers of these volumes.

2.3 Quantization errors

Both the camera CCD and the frame grabber have a limited resolution. This fact adds quantization errors to the images. There are many papers that discuss the effect of quantization in computer vision problems in general, see Kamgar-Parsi (1989) or Wong (1991) as an example or Olague Hernández (2005) in the context of corner detection. The effect of the quantization error in stereoscopic problems is studied as well in Rodríguez and Aggarwal 1990) and Sect. III of Blostein and Huang (1987). In our problem, the resolution is limited by the camera CCD. We have found that a limited resolution in the image is correlated with some human errors in the stereo part determination, thus the effect of the quantization will be incorporated in the quality index that will be discussed later, in Sect. 3.

2.4 Errors due to stereo pair determination

All equations in Sect. 2.1 assume that the coordinates of a point in the image planes of two cameras are known. But it may happen that the coordinates of this point can not be exactly measured: there may be a rounding error, due to the limited resolution of the camera, or the point may be

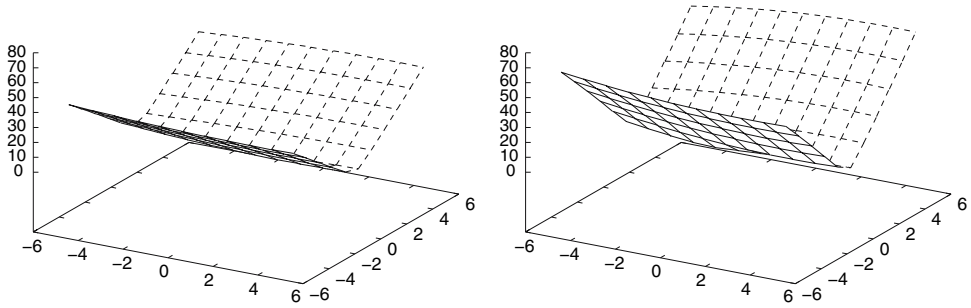

Fig. 6 Effect of the error made by the operator on the measurement error. Camera in the initial position (*left*) and with one meter displacement (*right*). The arrangement of the cameras affects the robustness of the measurement. The same error in a pair selection gives a greater error of the measurement if cameras are not arranged in the optimum positions

indistinguishable from its neighborhood. In these cases, an error may appear.

Applying Eq. (3) to both cameras (named "a" and "b") involved in the measurement of a dimension of interest, the system of Eq. (4) is obtained:

$$\begin{bmatrix} L_1^a - xL_9^a & L_2^a - xL_{10}^a & L_3^a - xL_{11}^a \\ L_5^a - xL_9^a & L_6^a - xL_{10}^a & L_7^a - xL_{11}^a \\ L_1^b - xL_9^b & L_2^b - xL_{10}^b & L_3^b - xL_{11}^b \\ L_5^b - xL_9^b & L_6^b - xL_{10}^b & L_7^b - xL_{11}^b \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} L_4^a - x^a + \delta x^a \\ L_8^a - y^a + \delta y^a \\ L_4^b - x^b + \delta x^b \\ L_8^b - y^b + \delta y^b \end{bmatrix} \tag{4}$$

The error in pair determinations is added to the bidimensional coordinates in the image acquired by cameras "a" and "b". If this error is zero, Eq. (4) represents two lines, as in Fig. 5, and their intersection gives us the tridimensional coordinates of the point. But, if this error is not zero, the result might be far from the true coordinates, depending on the angle between these lines.

As an example of the relevance of this kind of error, in Fig. 6 we have plotted the sum of the measurement errors in a vehicle, as a function of the inaccuracy in the matching of pairs made by the human operators. Axis x and y represent the number of pixels the operator fails in those directions. Axis z represents the modulus of the difference between the tridimensional coordinate obtained, and the real one. The left part of the figure shows the errors under the optimal camera positioning. In the right part, one of the cameras was displaced one meter to the right. As a consequence of this, the aggregate error increased from 47.2 to 70.7 cm. This means that, under identical human errors when selecting a stereo pair, the accuracy of some measurements could differ by more than 23 cm., depending on the arrangement of the cameras.

3 Fuzzy characterization of the measurement error

In the last section it was made clear that the positioning of the cameras was correlated with the average error, which in turn depended on the inaccuracy in the matching of pairs made by the operator. In this section we propose a method to characterize the error of the operator, so that we can assign a merit value to each arrangement of the cameras, and design an optimization algorithm that finds the optimal disposition.

3.1 Dispersion error and observation error

Let us suppose that an operator is asked to obtain the tridimensional coordinates of a given point, to annotate them, and then to repeat the same measurement again, a certain number of times. We do not expect to get a list with identical numbers, but a distribution of values with some dispersion, because not all the points he selects in the computer will be the same. Moreover, if a different operator reproduces the experiment, we expect a different distribution, because of different subjective perceptions of the images, skills using the pointing device, etc.

This dispersion in the values is one of the sources of uncertainty in our model. Let us assume it has a random nature, and that it depends on the human operator. For a worker named "A", his error in the selections can be represented by means of a bidimensional random variable X_A. If ω is an instance of the random experiment "select a pair of points", $X_A(\omega)$ is defined as a vector (e_x, e_y) that joins the point of the image that should have been selected with that point of the image that was actually selected. We will say that the probability distribution of X_A models the *dispersion error* of the operator "A".

A second source of error is a consequence of dealing with quantized images. Values (e_x, e_y) are not directly observable; we can only determine a small volume (not smaller than one pixel) that contains them. This is inherent to our procedure, because

- the cursor used to select the points on the screen has an appreciable area if compared to the tolerance in the measurement, and
- the image is digital, and pixels do not have null size (remember that low resolution surveillance cameras are being used).

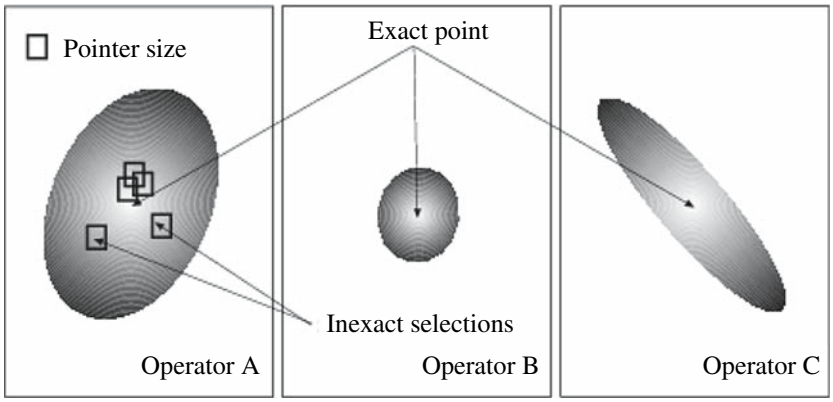

Fig. 7 Operators make errors when selecting the points for the stereo pair. If the probability of selecting one point, given the real position of the measurement, is known, a random set can be defined. The images of this random set are squares of the same size of the pointer used to select the point. The one point coverage function of this random set is taken as a fuzzy membership function, which models the typical error made by an operator. Dark zones in the figure mean "low membership"

Since this error prevents us from knowing the true coordinates selected, and following the usual nomenclature in fuzzy statistics, we will name it *observation error*.

3.2 Random sets and fuzzy based modeling

The observation error adds to the dispersion error mentioned before, but it is not of a random nature. The customary solution (Couso et al. 2004) consists in extending the random variable based characterization to a random sets based one, and to replace the variable X_A, with a random set Λ_A, that contains it:

$$X_A : X_A(\omega) \in \Lambda_A(\omega)\ \forall\omega. \tag{5}$$

That is to say, if the random variable X_A represents the difference between the true selection of the operator and the selection that the operator should have made, then the random set Λ_A represents the smallest square we can assure that contains this difference, provided that we do not know the exact point that the operator intended to select, but one pixel in the camera image that encloses the selection.

The one point coverage function of Λ_A will be understood as a fuzzy membership function (Goodman 1985). For a fuzzy set $\mathcal{X}_A$, this membership is

$$\mathcal{X}_A(x, y) = \mathrm{p}\{\omega : (x, y) \in \Lambda_A(\omega)\} \tag{6}$$

that joins dispersion and observation errors made by the operator "A" in his measurements.

If we knew that all measurements were to be carried out by operator "A", we could proceed to optimize the position of the cameras with respect to $\mathcal{X}_A$. But other operators will have different distributions of error (see Fig. 7). If the experiment is repeated with operators "B", and "C", etc. the fuzzy sets $\mathcal{X}_B$, $\mathcal{X}_C$ and so on will be obtained. We are interested in optimizing the worst case, therefore we define a compound model that aggregates all errors in the selection of a point in a set $\mathcal{X}$, defined as

$$\mathcal{X}(x, y) = \max(\mathcal{X}_A(x, y), \mathcal{X}_B(x, y), \ldots) \tag{7}$$

that will be estimated from a sample of the measurements made by every operator (Sánchez 1998; Cressie 1993).

3.3 Effect of the fuzzy error over the measurements

Once the error of a measurement has been characterized, it is necessary to study how the error made on a pair of measurements $\mathcal{X}$ and $\mathcal{X}'$, taken on two different projections of the same scene, affects the tridimensional measurement.

To propagate the fuzzy error, we will divide all fuzzy sets into their α-cuts, as shown schematically in Fig. 8. We can think of an α-cut of $\mathcal{X}$ as being an interval that covers the actual projection of the point being measured with a degree of truth $1-\alpha$. Therefore, the geometric place of all points in the space whose projections are compatible with this interval contains the actual point with truth $1-\alpha$. After having done the same with the set $\mathcal{X}'$, the intersection of both surfaces (the dashed polygon in Fig. 8) will contain the tridimensional coordinates with certainty $1-\alpha$. Note that every α-cut projects in a cone, and that the intersection of two of these cones is a tridimensional volume and not a planar surface as depicted in Fig. 8, which has been simplified.

It is easy to follow that volumes originated by nested cuts are also nested in a such way that when α varies from 0 to 1 a fuzzy set $\mathcal{C}$ defined over $\mathbf{R}^3$ it is obtained. The fuzzy set $\mathcal{C}$ models the point coordinates of the point that has been

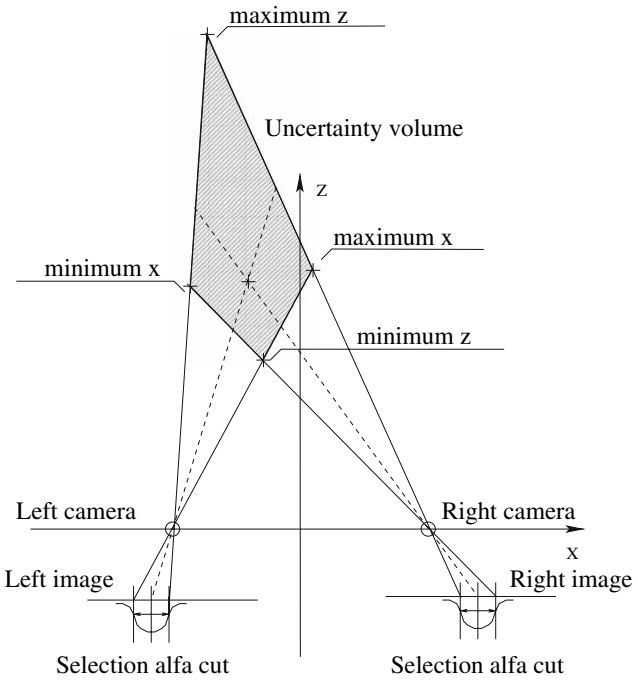

Fig. 8 Maximum and minimum x and z coordinates for a given α-cut of the input, for one dimension measurements, schematic determination. As can be seen, the closer the cameras are in x, the larger the uncertainty volume is in z

selected by the operator, as well as the imprecision associated to his selection.

3.4 Assignment of a merit value to the position of the cameras

Properties of the sets $\mathcal{C}$ vary as a function of the cameras' arrangement. If it is assumed that there is no bias in the measurements [this is, $E(X_i) = (0, 0)$ for every operator $i = A, B \ldots$] then the optimum arrangement of the cameras, for a given measurement, will be the one that makes the set $\mathcal{C}$ more specific. As the catalog of all standard measurements that can be performed on a vehicle is known, it makes sense to pose an optimization problem in which a given arrangement of the cameras will be related to the mean non-specificity of the sets $\mathcal{C}$ obtained when simulating all the possible measurements that will be taken.

In this work the definition of non-specificity used is that in Klir and Folger (1988),

$$U(\mathcal{C}) = \int \log ||\mathcal{C}_\alpha|| \, \mathrm{d}\alpha. \tag{8}$$

Therefore, the merit of an arrangement is the mean value of the estimation of Eq. (8) in the simulation of every measurement in the catalog, given a prototype vehicle.

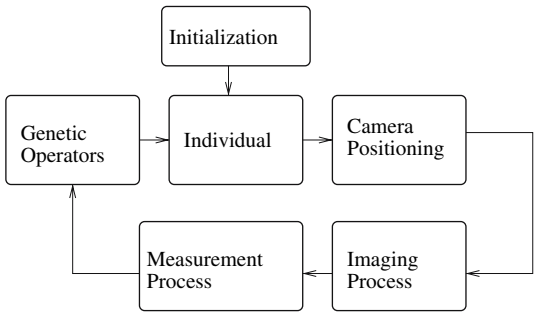

Fig. 9 Camera positioning optimization process

4 Optimization of the configuration of the cameras

Equation 8 assigns a numeric value to an arrangement of the cameras, for a given measurement over the vehicle. Consequently, constrained numerical optimization of this index over the set of all valid camera positions will produce the optimal disposition of the cameras for that measurement. A Genetic Algorithm will be used to perform this task. We found that GAs are well suited for this problem, as it is immediate to check that the function we wish to optimize is not continuous (consider, for instance, the corners of the plant), and it may have many different local minima (for example, it is clear that some different positions are equivalent between them, and they will produce the same objective value).

The whole optimization process that we propose is outlined in Fig. 9. In the following subsections, we will detail the representation, genetic operators and fitness function that will be used. Let us summarize for now the evaluation of an individual, that codifies a given arrangement of the cameras:

1. The process of the image acquisition is simulated, and synthetic images are generated. These images carry out information about the imperfections of the video sensors, as described in the preceding sections.
2. Using these synthetic images, the process of measurement, including a model of the human error, is recreated. A certain number of measurements are made, following the catalog of measurements in the standard procedure.
3. The result of Eq. 8, is evaluated for all of these measurements, and its mean value defines the fitness of the individual.

4.1 Representation

All the cameras used in the plant have the same lenses and the same CCD, so they only differ in their physical position and orientation. These are the only parameters that can be changed in the optimization process. In this way, each individual of the population will consist of the three angles and the three coordinates of the four cameras, that is, 24 real parameters.

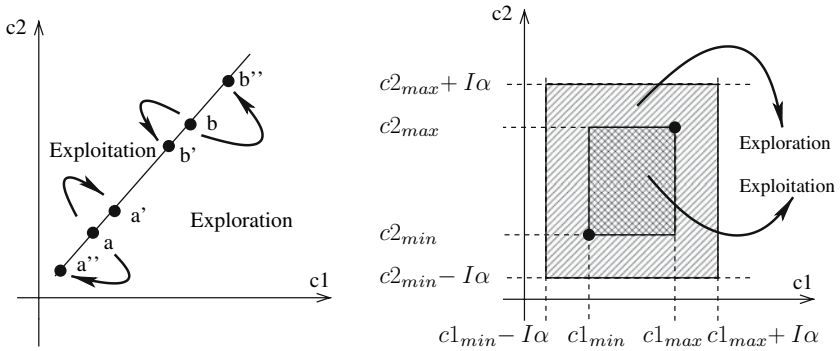

Fig. 10 Graphical interpretations of the crossover operators used in this work. *Top* arithmetic. *Bottom* BLX-α

Since each camera is mounted on a wall in the installation, one of the coordinates is fixed, and equals the distance from the coordinate origin to that wall, in x or y direction, depending on the orientation. That leaves us with 20 parameters, which have been codified with a standard floating-point based representation (Herrera and Lozano 1998).

4.2 Crossover

We have tested two different crossover operators well suited to the floating point-based representation: the arithmetic crossover (Michalewicz 1992) and the BLX-α crossover (Eshelman and Schaffer 1993).

Arithmetic crossover between two individuals a and b is defined as follows: first, a coefficient α is randomly generated such that $\alpha \in (-c, c)$ (in the experiments performed we take $c = 0.25$). Second, we emit the offspring: $a - \alpha * (a - b)$ and $b + \alpha * (a - b)$. In the left part of Fig. 10, we show graphically how this crossover operator works, with individuals comprising two genes. Note that the results are a subset of all possible lineal combinations of the original individuals. Because of this, exploration and exploitation is reduced to the straight line defined by the individuals.

By contrast, BLX-α crossover pays more attention to the exploration of the space of solutions. For example, let two chromosomes, comprising n genes each, be $C_1 = (c_1^1, c_2^1, \ldots, c_n^1)$ and $C_2 = (c_1^2, c_2^2, \ldots, c_n^2)$. The offspring of the BLX-α crossover comprises $H_k = (h_1^1, h_2^1, \ldots, h_n^1)$ and $H_k = (h_1^2, h_2^2, \ldots, h_n^2)$, where h_i^k, with $i = 1 \ldots n$ and $k = 1, 2$, are randomly chosen from $[c_{\min} - I\alpha, c_{\max} + I\alpha]$, and where $c_{\max} = \max\{c_i^1, c_i^2\}$, $c_{\min} = \min\{c_i^1, c_i^2\}$ and $I = c_{\max} - c_{\min}$.

In the right part of Fig. 10 it is shown how this operator works. Note that, in this case, exploration and exploitation are not constrained to the straight line defined by the individuals, but they are extended to the rectangular areas shaded with a different pattern in the same figure. Our experiments do not contradict results in Herrera (2002), and we also found that the BLX-α crossover operator was able to discover more efficient solutions to our problem than the arithmetic crossover.

4.3 Mutation

In this work we use the mutation operator proposed in Michalewicz (1992), because it performs a uniform search in the space of solutions at the beginning of the execution and becomes more local as the algorithm goes on. Let T be the total number of generations and t the actual generation. Given a chromosome of generation t, $C_t = (c_1, c_2, \ldots, c_k, \ldots, c_n)$, where one of its genes c_k, with range in $[c_{ki}, c_{kd}]$, is selected to be mutated, the chromosome obtained after the mutation is $c^{t+1} = (c_1, \ldots, c'_k, \ldots c_n)$, where

$$c'_k = \begin{cases} c_k + \Delta(t, c_{kd} - c_k) & \text{if } p = 0 \\ c_k - \Delta(t, c_k - c_{ki}) & \text{if } p = 1 \end{cases} \tag{9}$$

where p is a random number from 0, 1 and

$$\Delta(t, y) = y\left(1 - r^{(1-\frac{t}{T})^b}\right) \tag{10}$$

gives a value in the interval $[0, y]$, so that the probability that $\Delta(t, y)$ is near to zero increases with t.

In Eq. 10, r is a random number in the interval [0, 1] and b is a number chosen by the user that specifies the degree of dependency between that probability and the number of generations.

4.4 Fitness

The fitness function is oriented to solve a constrained optimization problem. Observe that (a) the average non-specificity of the error should be minimal, and (b) the positions of the camera should be feasible. The second criterion, in turn, splits in two: (1) the whole volume of the vehicle must be visible from all cameras and (2) the camera must be on a wall or column, never on a door.

There is an obvious ranking of objectives, thus we do not need to resort to multiobjective genetic algorithms. Since we will be using a steady state GA, with tournament based selection, it is only necessary to define a lexicographic ordering:

one impossible arrangement is always worse than a valid one, and solutions where some measurements are not possible are worse than solutions where every part of the vehicle is visible.

4.4.1 Average non-specificity

Calculation of the first fitness component is done by simulating the determination of the tridimensional coordinates of the control points. First, using a projective model, the images of the pair of cameras needed to take the measurement are obtained. Then, the bidimensional coordinates of every control point in the image are distorted, by adding to them the fuzzy set that models the error made by the operator. Then, for three α-cuts (0.25, 0.50 and 0.75) the maximum and minimum values for the tridimensional coordinates are obtained (see Fig. 8) and the non-specificity of the fuzzy extension of the measurement is estimated. The final value is the mean of these non-specificities for every point of control. In our experiments it was supposed that X_i are gaussians, and that the side of the square needed to define Λ_i is known.

4.4.2 Visibility of the control volume

The second component of the fitness measures the percentage of the control volume visible from the cameras. To estimate it, we count the number of hidden measurement points in the control volume. Each hidden point contributes to the error with an amount proportional to its distance to the visible measurement point closer to it.

4.4.3 Valid positions of the cameras

The last component of the fitness is null when the camera can be physically attached to the wall at the given position, and depends on the distance to the nearest feasible mount otherwise. In practice, this situation occurs when the coordinates are on a door (the distance to its frame would be computed) or when the coordinates are outside the plant (the distance to the center of the nearest wall would be evaluated).

4.5 Numerical results

In Fig. 11, a schematic representation of the plant is shown, including the most relevant magnitudes and the places where cameras can be attached.

The error of the human operators has been modeled by means of fuzzy memberships corresponding to coverage functions of random sets whose images are squares, with a side measuring two pixels, and whose centers follow a bidimensional normal distribution with covariance matrix $4 \cdot I$. The measurement pattern is a prism of $1.5 \times 1.5 \times 3.0$ m (see Fig. 1).

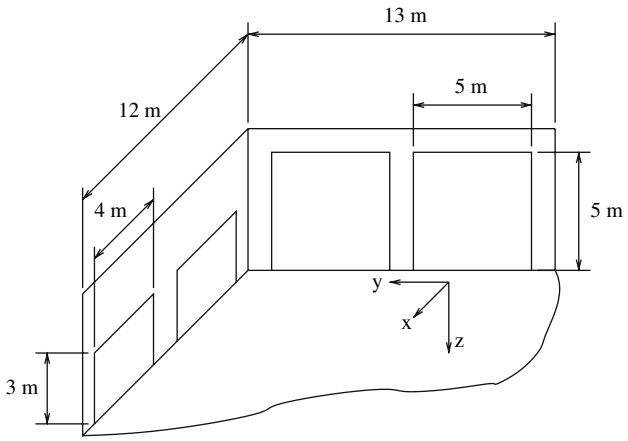

Fig. 11 Simplified model of the plant, used in the experiments. The axis names and the set of feasible positions for camera mountings are shown

Table 1 Genetic algorithm parameters

Population size	100
Tournament size	5
Mutation probability	0.01
Number of generations	5000

Table 2 Coordinates of the cameras installed in the initial physical position, selected using the expert's criteria

Camera	x	y	z
1	−600	1000	−250
2	−100	1100	−75
3	100	1100	−75
4	600	1000	−250

The parameters of the genetic algorithm are shown in Table 1, the stopping criteria was a maximum number of generations. Experiments with arithmetic and blx-α crossover were performed.

If the cameras are initially installed in the positions shown in Table 2, with the angles shown in Table 3, the mean error of the system is 14 cm. When the configuration is optimized following the approach presented in this paper, the error falls to 6.5 cm. using arithmetic crossover and 5 cm using blx-α crossover. The coordinates of the physical positions of the cameras using arithmetic crossover are shown in Table 4, with the angles shown in Table 5. The coordinates of the physical positions of the cameras using arithmetic crossover are shown in Table 6, with the angles shown in Table 7.

In Fig. 12, the original placement, the optimized one using arithmetic crossover and the optimized one using blx-α, are shown. The optimized physical position of a camera obtained using arithmetic crossover is represented by numbers surrounded by a square, where the number denotes the number

Table 3 Camera angles for camera positions given in Table 2

Camera	α	β	γ
1	0.19635	−1.5708	−1.2566
2	0	1.5708	−1.2566
3	0	1.5708	−1.2566
4	0.3927	1.5708	−1.2566

Table 4 Coordinates of the cameras installed in the optimized physical positions following the proposed approach and arithmetic crossover

Camera	x	y	z
1	−600	−117.56	−107.08
2	587.1	1100	−2.47
3	−42.13	1100	−83.14
4	600	−163.08	77.94

Table 5 Camera angles for camera positions given in Table 4

Camera	α	β	γ
1	−0.42376	0.498069	−0.181525
2	0.265457	−0.0755015	−0.0166876
3	−1.45492	2.74395	0.815273
4	0.012051	2.09327	−0.220714

Table 6 Coordinates of the cameras installed in the optimized physical positions following the proposed approach and blx-α crossover

Camera	x	y	z
1	−600	−127.97	−78.34
2	−270.96	1100	−262.75
3	583.4	1100	278.41
4	600	−114.35	−109.35

Table 7 Camera angles for camera positions given in Table 6

Camera	α	β	γ
1	0.4472	−0.4350	−0.4999
2	−0.5279	−0.9901	0.2864
3	1.8574	0.8398	0.4538
4	0.7186	0.0192	−1.5429

of the camera, as shown in Table 4. The optimized physical position of a camera obtained using blx-α crossover is represented by numbers surrounded by a circle, where the number denotes the number of the camera, as shown in Table 6. Cameras installed in the places suggested by the expert manner are represented by numbers surrounded by a hexagon, and the numbering matches that shown in Table 2.

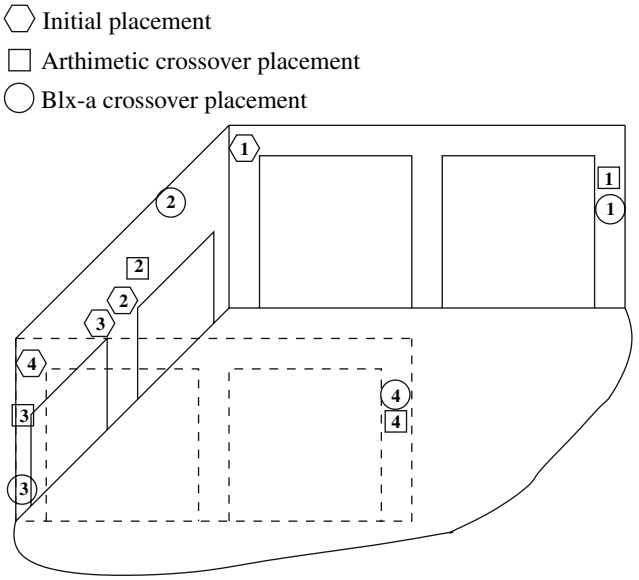

Fig. 12 Schematic representation of the physical positions of the cameras, original and optimized

It must be highlighted that the cameras are rearranged in a more uniform manner around the plant, and this was an expected result; but they are also arranged in height: this was not obvious. Once the optimization process has revealed it, it also seems a logical conclusion: when we first placed cameras by hand, we intuitively tried to separate them as much as possible, but we positioned all of them at a constant height. The genetic search discovered that this separation could have been higher if their heights were different.

Another interesting aspect of the camera placement obtained by the algorithm is that it is such that the object being measured appears in an orientation so that the longest dimension is aligned with the image axis with more resolution, as can be seen in Fig. 13. Again, it is obvious that this decreases the measurement error, since the resolution increases in that direction.

As stated before, two types of crossover were tested. In Fig. 14, the corresponding boxplots of 30 runs of the algorithm using blx-α (left) and arithmetic crossover (right) is shown. As can be seen, there is a significant difference between both types of crossover, and the obvious choice is blx-α.

Finally, the stochastic nature of the genetic algorithm is shown in Fig. 15. In this figure the positions of the cameras obtained from 500 runs of the algorithm are shown. As can be seen, there is some degree of dispersion of the results, sometimes the positions are not valid being positioned on the doors of the installation (represented as squares) but there is also some aggregation of camera positions in some places of the installation that agree with the preceding comments: the optimal placement of the cameras is such that they are

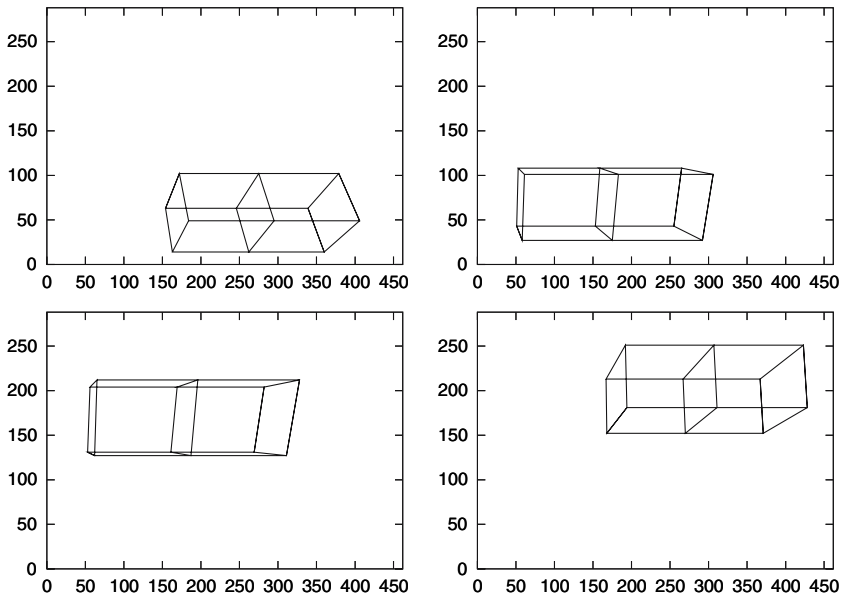

Fig. 13 Synthetic images of the measurement pattern obtained with the cameras placed as suggested by the algorithm. Observe that the orientation of the major axis is parallel to the image axis with more resolution

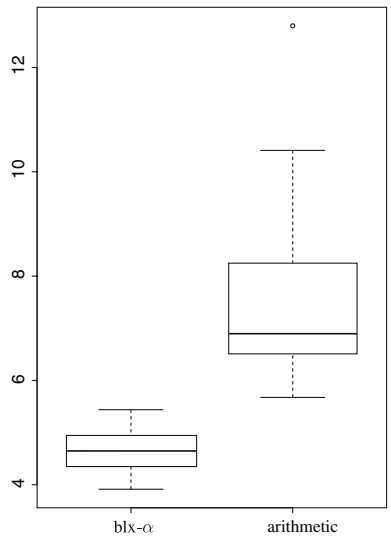

Fig. 14 Boxplot of measurement error obtained from 30 runs of the algorithm using blx-α and arithmetic crossover. As can be seen, there are significant differences in the results and blx-α seems to be the one with less error

at different heights and spread around the plant as much as possible.

5 Concluding remarks and future work

In the work described in this paper, fuzzy sets have been used to model an experiment in which two sources of imprecision are present: one of them is due to the randomness of the human operator behavior. The other depends on the inaccuracy in the observation of the results of the experiment.

Fuzzy sets were used to design a model of either source of error. Every selection in an image was associated a fuzzy set, that assigned a decreasing confidence to its neighborhood. When determining the spatial coordinates of one point, the information provided by the fuzzy sets was used to build a new fuzzy set associated to them. This last fuzzy set is more specific when the camera is correctly positioned, thus a genetic algorithm was used to find the positions for which the specificities are maximum, assuring that the positions of all

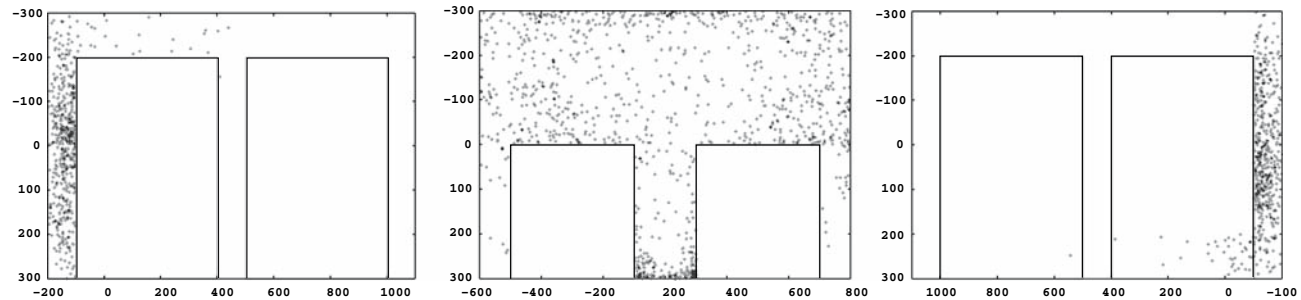

Fig. 15 Different camera locations obtained from 500 runs of the algorithm. The stochastic nature of the GA is shown, along with some aggregation of solutions in some places of the installation

cameras are valid, and the control volume can be seen from all of them. This produces a measurement system with increased robustness against user errors. The optimized placement gave a measurement system with a valid average tolerance, while using cheap video equipment that in principle was not designed to do this task.

In future works, we intend to improve the accuracy of the measurements by means of a revamped interface. In the new system, when the operator has finished selecting the points that are to be matched, geometric transformations will be applied to the images so that the operator's selection can be fine tuned by the computer. The tuning will consist in genetically maximizing the local matching between the set of transformed images in the neighborhood of these points (Santamaría et al. 2007). Afterward, the results will be presented to the operator so that he validates the tuned points (or decides to reject the changes). The main problem with this approach is the time needed to find the best matching. In this sense, we find promising a new definition of the fuzzy similarity defined in Tolt and Kalaykov (2006), which we intend to combine with the genetic search in order to reduce the elapsed time between the initial selection of the points and their validation by the operator. We are also interested in knowing whether the best position of the cameras in the improved system coincides with that of the former system. Finally an improvement of the approach presented here consists in the combination of the model in Olague and Mohr (2002) with the fuzzy analysis presented here.

References

Abdel-Aziz Y, Karara H (1971) Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. In: Proceedings of the ASP Symposium on Close Range Photogrammetry. Fall Church, VA: American Society of Photogrammetry, pp 1–18

Bhat DN, Nayar SK (1998) Ordinal measures for image correspondence. IEEE Trans Pattern Anal Mach Intell 20:415–423

Blostein SD, Huang TS (1987) Error analysis in stereo determination of 3D point positions. IEEE Trans Pattern Anal Mach Intell 9: 752–765

Brown MZ, Burschka D, Hager GD (2003) Advances in computational stereo. IEEE Trans Pattern Anal Mach Intell 25:993–1008

Chen L, Armstrong CW, Raftopoulos DD (1994) An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique. J Biomech 27:493–500

Couso I, Sánchez L, Gil P (2004) Imprecise distribution functions associated to a random set. Inf Sci 159:109–123

Cressie NAC (1993) Statistics for Spatial Data. Wiley, London

Duda R, Hart P (1973) Pattern classification and scene analysis. Wiley, New York

Eshelman LJ, Schaffer JD (1993) Real-coded genetic algorithms and interval schemata. In: Foundations of genetic algorithms.2. Morgan Kaufmann Publishers, Inc., San Mateo, pp 187–202

Faugeras O, Hotz B, Matthieu H, Vieville T, Zhang Z, Fua P, Theron E, Moll L, Berry G, Vuillemin J, Bertin P, Proy C (1993) Real time correlation-based stereo: algorithm, implementations and applications INRIA Technical Report 2013

Faugeras O, Luong QT (2001) The Geometry of Multiple Images. The MIT Press, Cambridge

Goodman N (1985) Uncertainty Models for Knowledge-based Systems. North-Holland, Amsterdam

Hartley R, Zisserman A (2000) Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge

Hatze H (1988) High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT-approach. J Biomech 21:533–538

Herrera F, Lozano M (1998) Tackling real-coded genetic algorithms: Operators and tools for the behavioural analysis. Artif Intell Rev 12:265–319

Herrera F, Lozano M, Pérez E, Sánchez AM, Villar P (2002) Multiple crossover per couple with selection of the two best offspring: an experimental study with the BLX-α crossover operator for real-coded genetic algorithms. In: Advances in artificial intelligence—IBERAMIA 2002: 8th Ibero–American Conference on AI LNCS 2527/2002, pp 392–401

Hinrichs RN, McLean SP (1995) NLT and extrapolated DLT: 3D cinematography alternatives for enlarging the volume of calibration. J Biomech 28:1219–1224

Hoff W, Ahuja N (1989) Surfaces from stereo: integrating feature matching, disparity estimation, and contour detection. IEEE Trans Pattern Anal Mach Intell 11:121–136

Kamgar-Parsi B (1989) Evaluation of quantization error in computer vision. IEEE Trans Pattern Anal Mach Intell 11:929–940

Klir G, Folger T (1988) Fuzzy sets, uncertainty, and information. Prentice Hall

Kwon YH, Lindley S (2000) Applicability of 4 localized-calibration methods on underwater motion analysis. In: XVIII International Symposium on Biomechanics in Sports

Marzan GT, Karara HM (1975) A computer program for direct linear transformation solution of the collinearity condition, and some applications of it. In: Proceedings of the symposium on close-range photogrammetric systems. Falls Church, VA: American Society of Photogrammetry, pp 87:109

Mason SO, Grünt A (1995) Automatic sensor placement for accurate dimensional inspection. Comput Vis Image Underst 61:454–467

McVey ES, Lee JW (1982) Some accuracy and resolution aspects of computer vision distance measurements. IEEE Trans Pattern Anal Mach Intell 4:646–649

Michalewicz Z (1992) Genetic algorithms + data structures = evolution programs. Springer, Heidelberg

Olague G, Mohr R (2002) Optimal Camera Placement for Accurate Reconstruction. Pattern Recognit 35:927–944

Olague G (2002) Automated photogrammetric network design using genetic algorithms. Photogramm Eng Remote Sens 68:423–431

Olague G, Hernández B (2005) A new accurate and flexible model-based multi-corner detector for measurement and recognition. Patter Recognit Lett 26:27–41

Olague G, Dunn E (2006) Development of a practical photogrammetric network design using evolutionary computing. Photogramm Rec 117:22–38

Rodríguez JJ, Aggarwal JK (1990) Stochastic analysis of stereo quantization error. IEEE Trans Pattern Anal Mach Intell 12:467–470

Sánchez L (1998) A random sets-based method for identifying fuzzy models. Fuzzy Sets Syst 3:343–354

Santamaría J, Cordón O, Damas S, Alemán I, Botella M (2007) A scatter search-based technique for pair-wise 3D range image registration in forensic anthropology. Soft Comput 9:819–828

Scharstein D, Szeliski R (1998) Stereo matching with non-linear diffusion. Int'l J Comput Vis 28:155–174

Shapiro R (1978) Direct linear transformation method for three-dimensional cinematography. Res Quart 49:197–205

Tolt G, Kalaykov I (2006) Measures based on fuzzy similarity for stereo matching of color images. Soft Comput 10:1117–1126

Tomasi C, Manduchi R (1998) Stereo matching as a nearest-neighbor problem. IEEE Trans Pattern Anal Mach Intell 20:333–340

Verri A, Torre V (1986) Absolute depth estimates in stereopsis. J Opt Soc Am 3:297–299

Wong PW (1991) On quantization errors in computer vision. IEEE Trans Pattern Anal Mach Intell 13:951–956