Evolutionary-based 3D reconstruction using an uncalibrated stereovision system: application of building a panoramic object view

**Alain Koch · Albert Dipanda ·
Claire Bourgeois République**

Abstract In this paper, we propose an original evolutionary-based method for 3D panoramic reconstruction from an uncalibrated stereovision system (USS). The USS is composed of five cameras located on an arc of a circle around the object to be analyzed. The main originality of this work concerns the process of the calculation of the 3D information. Actually, with our method, 3D coordinates are directly obtained without any prior estimation of the fundamental matrix. The method operates in two steps. Firstly, points of interest are detected in pairs of images acquired by two consecutive cameras of the USS are matched. And secondly, using evolutionary algorithms, we jointly compute the transformed matrix between the two images and the respective depth of the points of interest. The accuracy of the proposed method is validated through a comparison with the depth values obtained using a traditional method. In order to perform 3D panoramic object reconstruction, the process is repeated for all the pairs of consecutive cameras. The 3D points thus obtained throughout the successive steps of the process which correspond to the different points of interest, are combined in order to obtain a set of 3D points all around the analyzed object.

Keywords Panoramic 3D reconstruction · Evolutionary algorithm · Uncalibrated system

1 Introduction

Panoramic 3D reconstruction is a recent area of research which began about 10 years ago. It is used in various domains such as medicine [32], security [30], virtual reality or robotics

A. Koch · A. Dipanda · C. Bourgeois République (✉)
LE2I, Faculté des Sciences, Université de Bourgogne,
Aile de l'ingénieur 9, Avenue Alain Savary, BP 47870, 21000 Dijon, France
e-mail: claire.bourgeois-republique@u-bourgogne.fr

A. Koch
e-mail: alain.koch@u-bourgogne.fr

A. Dipanda
e-mail: adipanda@u-bourgogne.fr

[25]. The goal of panoramic 3D reconstruction is to achieve a global internal representation of a scene or a surrounding view of an object. 3D panoramic reconstruction methods can be divided into two broad classes of system. The first class of methods uses only one camera while the second one uses several cameras. The double lobed mirror [11] system for example, belongs to the first class. It includes a double hemisphere mirror with different diameters and the camera is located above the double mirror. The panoramic 3D reconstruction is processed using the images obtained from the two mirrors, which provide two different points of view of the scene. In the same class, we can also include the cylindrical camera acquisition method [5]. It is based on 2D panoramic images of a scene. In this case, a panoramic 3D reconstruction requires at least five images from different locations in the scene. This kind of systems often uses a mirror [11, 27] which can distort the scene, and interfere with the detection of some basic geometrical forms. A panoramic 3D reconstruction can also be obtained with a single camera by moving the camera around the scene [26, 28], which allows different points of view to be taken. The respective locations or the movement of the camera have to be exactly known which makes the approach comparable to the second class of methods in which the panoramic stereovision systems are composed of two or more cameras. In this second class, more data are required to calculate the reconstruction. Moreover, the images taken by different cameras need to be matched [23] before the reconstruction calculation. The methods in this class can further be classified into two groups: The calibrated-based methods [10, 33] and the uncalibrated systems. The calibrated methods are characterized by an initialization step called the system calibration. The goal is to obtain the matrix of correspondence between the real coordinate system and the image coordinate systems. Extrinsic and intrinsic parameters are needed to achieve this aim. Some uncalibrated systems use object movement in the scene [7, 14, 16] to obtain the 3D panoramic reconstruction. Usually, six predefined movements, two per axis (x, y and z), are executed in order to calculate the transformation between the real and observed movements. Three transformations are calculated respectively for the different axes, which allows the calculation of the global transformation between the real world system and each of the camera system. By contrast, uncalibrated systems use Points Of Interest (POIs) obtained on different images in order to calculate the transformation between the different cameras. These POIs can be determined either with an object the geometry of which is known, for example a grid [24], or by placing some markers on the studied object. Other systems use the properties of a specific type of object, such as the size, the axis of symmetry or the shape, to extract from the image one or more parameters, which allows the calibration (sometimes only partially) of the system. For example, the shape of a hand, especially the edge of fingers, is used to reconstruct and to recognize the hand gesture [31].

In this study, our ultimate goal is to make a 3D panoramic reconstruction of an object without involving the camera parameters. The proposed 3D panoramic reconstruction method is composed of three steps. Firstly, we detect POIs on different sides of the object. Secondly, without any calibration, we calculate the 3D coordinates of the POIs. Finally, the 3D coordinates of the POIs of the different sides are matched using a 2D surrounding view of the object, in order to make a 3D panoramic reconstruction. We use an uncalibrated stereovision system (USS) made up of five cameras which are circularly located around the studied object. Five images of the object are then acquired with the respective cameras. Markers are placed on the object in order to facilitate the detection of POIs. The POIs detected on the images acquired by consecutive cameras are matched. Since the system is uncalibrated, some (or all the) parameters that allow the calculation of the 3D transformation between a pair of images are missed, which leads to a very hard non-

linear problem. This problem has been addressed [19] using Genetic Algorithms (GAs). In this article, the authors use GAs to calculate the projection matrix between two uncalibrated images. In contrast, we propose to directly calculate the 3D information without prior knowledge of the projection matrix. We have expressed this problem as a minimization of a global function that we propose to perform by using Evolutionary Algorithms (EAs). To our knowledge, our work is currently the only work based on EAs which addresses this problem in this direction. EAs are a stochastic search method which is inspired by natural selection and the principles of evolution [12, 13]. EAs have received a great deal of attention in the recent past and are widely used in diverse areas of image processing [3, 6, 17, 18, 21, 22].

This paper is organized as follows: In section 2, we briefly describe the background of epipolar geometry which is used for the calibration methods. In section 3, we describe the USS and detail the two steps of the method: the 3D partial reconstruction and the panoramic reconstruction methods. Section 4 is devoted to the experimental results. Finally, in section 5, after concluding our work, we mention some of the further tasks that we must deal with.

2 Background of epipolar geometry

Usually a 3D reconstruction process requires a stereovision system composed of two or more cameras (see Fig. 1). The 3D coordinates are calculated by triangulation which implies that the camera locations in the scene and the coordinates of the points in each image must be known. The pinhole camera model is usually used to calibrate the cameras. For this purpose, the intrinsic and the extrinsic parameters must be calculated [9, 10]. The extrinsic parameters allow the calculation of the transformation between a real point in the world system and its corresponding point in the camera system, while the transformation between the image plane and the camera system is provided by the intrinsic parameters. The intrinsic and extrinsic parameters must be calculated for each camera.
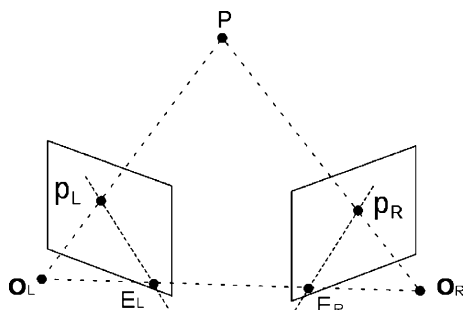
For a given camera the intrinsic parameter matrix Ic is obtained as follows:

$$[p] = [Ic][P]$$

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} \alpha u & 0 & u0 & 0 \\ 0 & \alpha v & v0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ s \end{bmatrix} \tag{1}$$

where (u, v) are the coordinates of 2D point p in the image plane, (X, Y, Z) are the coordinates of 3D point P on the object in the camera system, s is a scale factor

Fig. 1 Theoretical stereovision acquisition system

(generally, equal to one) and (u_0, v_0) are the coordinates of the optical centre in the image plane.

The values of the parameters α_u and α_v are calculated as follows:

$$\alpha_u = k_u f \tag{2}$$

$$\alpha_v = -k_v f \tag{3}$$

Where k_u (resp. k_v) is a vertical (resp. horizontal) scale factor of the camera and f is the distance named the focal length, α_u, α_v, u_0 and v_0 are the four intrinsic parameters in matrix Ic.

The extrinsic parameters are calculated using the following formula:

$$[A] = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

Matrix A includes a rotation R and a translation T between the world and the camera systems. These elements are known. From matrices I_c and A we obtain matrix M which allows the calculation of the 2D coordinates of a world point in the image. Matrix M is given by:

$$[M] = [Ic][A] = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \tag{5}$$

$$p = [M]P \tag{6}$$

In order to obtain the respective left and right camera matrices M and M', the same test pattern is used. For each camera, at least six points must be detected in the image to solve the previous equation system. Then with A and A' the extrinsic matrix, A_s can be calculated as follows:

$$[A_s] = [A'][A]^{-1} = \begin{bmatrix} r_{11} & r_{21} & r_{13} & b_x \\ r_{21} & r_{21} & r_{23} & b_y \\ r_{31} & r_{21} & r_{33} & b_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} r_1 & b_x \\ r_2 & b_y \\ r_3 & b_z \\ 0 & 1 \end{bmatrix} \tag{7}$$

Let us consider a given world point P viewed respectively as a point $p_{Lc}(x_{Lc}, y_{Lc}, z_{Lc})$ in the left camera system and $p_{Rc}(x_{Rc}, y_{Rc}, z_{Rc})$ in the right one, Eq. 8 allows the two coordinate systems to be aligned:

$$p_{Rc} = [A_s]p_{Lc} \tag{8}$$

3D point P is at the intersection of the two lines defined by the projection centre of the camera (O_L for the left camera and O_R for the right one) and point $p_L(u_L, v_L)$ in the left image, respectively point $p_R(u_R, v_R)$ in the right image.

To match a point p_L corresponding in the left image system with the point p_R in the right image system, Eq. 9 is used to define the epipolar line which corresponds to the intersection between the right image plane and the plane defined by the triplet (P, O_L, O_R):

$$(b_z r_2 \bullet p - b_y r_3 \bullet p)u_R + (b_x r_3 \bullet p + b_z r_1 \bullet p)v_R = b_x r_2 \bullet p_L - b_y r_1 \bullet p_L. \tag{9}$$

The problem at hand becomes how to determine the exact location of point p_R on the epipolar line. This can be performed for example by using the well-known block matching method [2].

The equation system to calculate the coordinates of point P is given by:

$$u_R = \frac{(r_{11}u_L + r_{12}v_L + r_{13})Z + b_x}{(r_{31}u_L + r_{32}v_L + r_{33})Z + b_z} \tag{10a}$$

$$X = u_L Z \tag{10b}$$

$$Y = v_L Z \tag{10c}$$

In Fig. 1, Point E_L (respectively E_R) is named the left epipole (respectively the right epipole) and it is the projection of the focal point of the right camera O_R (resp. focal point of the right camera O_L) into the left image system (resp. right image system). All the epipolar lines in the same image go through the epipole.

3 Proposed 3D panoramic reconstruction method

In this section, we describe the proposed 3D panoramic reconstruction method. We introduce our image acquisition system followed by the reconstruction method based on Evolutionary Algorithms. Then, we describe the panoramic 3D reconstruction process.

3.1 Description of our Uncalibrated Stereovision System (USS)

Our USS is composed of five web cameras which are positioned circularly around the object to be analyzed (see Fig. 2). Two consecutive cameras are separated by an angle of 30°. The partial 3D reconstruction (i.e. the 3D reconstruction of the part of the object that is viewed by two consecutive cameras) is performed thanks to the images acquired by this pair of cameras. The 3D reconstruction process requires matching POIs obtained from the two images. Since we consider that the object to be studied may have no texture, the standard methods [1, 4, 8, 15] usually used to detect POIs are not efficient in this case and can lead to wrong results. Thus, some markers are placed on the object.

3.2 Depth calculation using evolutionary algorithm

Our preliminary goal is to achieve a partial object 3D reconstruction based on two images. For this purpose, we propose the following two steps: firstly, we detect the pairs of POIs which are viewed on both images. Secondly, we determine the respective depth values of these POIs using an Evolutionary Algorithm.

Fig. 2 A view of our acquisition system. The analysed object in this image is a yellow car
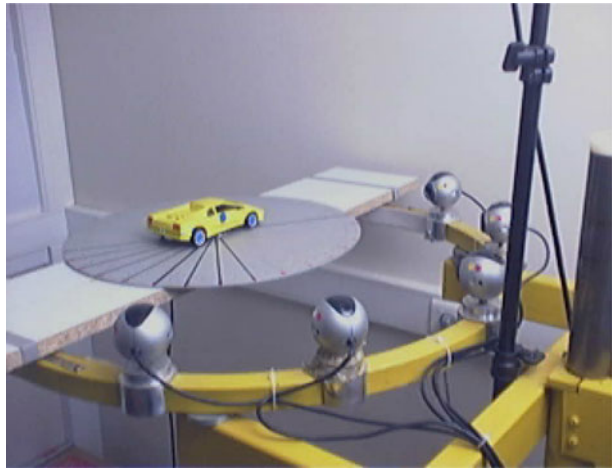
The main idea we are putting forwards is the following: two corresponding POIs which are respectively detected in both images represent the same physical 3D point as illustrated on Fig. 1. Hence, in order to transform a POI in one image into its matched point in the second image, we have to determine the depth of this physical 3D point.

3.2.1 Detection and matching of POIs

The main step on which the proposed method relies on is the detection of POIs on both images acquired by two consecutive cameras of the USS. These POIs are then matched. As mentioned previously, markers are placed on the object to be analyzed in order to facilitate the determination of the POIs.

The POI matching process is as follows: the markers are respectively detected in both images by using a color threshold method [30]. It performs a color segmentation based on the hue and the saturation components of each image. This method allows strong color sensitivity and a small sensitivity to the luminosity variation between two images or two scenes.

The different colors of the marker must be known before the use of the method. The different steps of the POI detection are the following:

1. Choose a marker color.
2. Change the RGB color system in the HSL.
3. Determine the threshold of the marker color according to the hue and the saturation values, which provides a binary image, (i.e. all the detected marker points are in white, while the remainder of the image is in black).
4. Scan the image in order to detect all the connected white points. When a *new non-connected* white pixel is found, a new marker is added in a list. At the end of this step, all the markers of a given color are detected.
5. Eliminate the possible noise using a criterion based on the size of the markers.
6. Determine the highest and the lowest points of each marker (i.e. the POIs).
7. Choose another marker color and go to two until all the colors are processed.

It must be noted that, for a given marker, we have chosen to use the highest and the lowest points as POI since they are less dependent on the location of the camera during the image acquisition. And this choice increases the efficiency of the Evolutionary algorithms

by increasing the precision of the reconstruction. With two points per marker, the distortion and the change of scale is detected immediately.

A pair of sets of POI obtained from both images is then constituted. Finally, these points are matched between the two images by executing the following steps:

1. Count the markers of each color in each image. If an image has fewer markers than the other, the marker in excess is detected because it must be on the far right or left of the object. This marker is then taken out from the marker list.
2. Create all the possible matches between the remaining markers with the same color.
3. Eliminate all the impossible cases: permutations of two or more markers between images.
4. Keep only the correct matching.

3.2.2 Evolutionary algorithm

Evolutionary Algorithms are adaptive procedures that find solutions to problems by using an evolutionary process based on natural selection. An EA uses a finite population of potential solutions to a problem. Each individual solution is encoded as a chromosome made up of a string of genes which take values in either a binary or a non-binary alphabet. An EA comprises three main stages: evaluation, selection and mating. They are applied cyclically and iteratively until saturation or other boundary condition is satisfied. At the evaluation stage, each chromosome is assigned a fitness value which represents its ability to solve the problem. At the selection stage, chromosomes are chosen, based on their fitness in such a way that the better chromosome is more likely to be selected. At the mating stage, crossover and mutation operations are performed. The crossover operation recombines pairs of selected chromosomes, also called parent chromosomes, to form two new offspring. The mutation operation creates new offspring by modifying one or more genes of a chromosome chosen randomly from the mating pool. From generation to generation, this process leads to increasingly better chromosomes and to near-optimal solutions.

3.2.3 Chromosome encoding

The input of our reconstruction problem is made up of a set of POIs which are the highest and lowest points of the markers. We state that for a given marker, these two points have the same depth value. So in what follows, we consider that the depth marker equals the depth of one of its two corresponding POIs. A chromosome of our EA represents the respective depth values of the markers. Thus, a chromosome is composed of N genes, where N is the marker number. And, a gene is a real number which is comprised between 0 and the maximum length of the object to be analyzed. Figure 3 illustrates an example of a chromosome with five genes (i.e. the marker depth values).

3.2.4 Fitness function

Our goal is to find the depth values of the different N markers detected on the image. It is important to mention that the known parameters in our USS are only the rotation angles between the cameras and the focal lengths of the cameras. In order to evaluate the N marker depth values, we calculate the global transformation which projects all the POIs from the left image into their corresponding matched POIs on the right image. This transformation

Fig. 3 Chromosome encoding

Name of the markers	M1	M2	M3	M4	M5
Depth values of the markers	X_{M1}	X_{M2}	X_{M3}	X_{M4}	X_{M5}
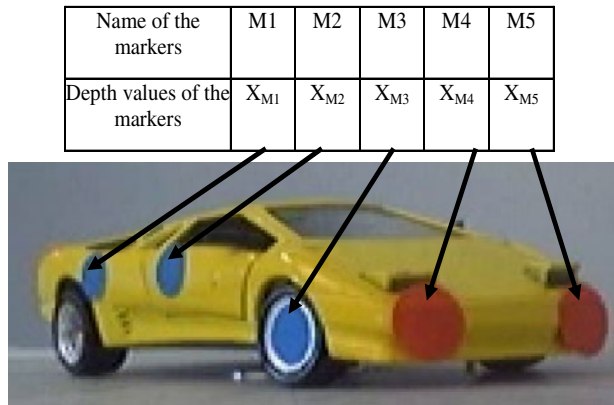
requires four steps. Let us consider the couple of points $Pl_i(Xl_i, Yl_i)$ and $Pr_i(Xr_i, Yr_i)$ which represent the ith POI in the left image and its corresponding point in the right image. Point $Pl_i(Xl_i, Yl_i)$ will be transformed as follows:

1. A perspective projection T_1 is achieved on $Pl_i(Xl_i, Yl_i)$, which gives the 3D point $Plc_i(Xlc_i, Ylc_i, Zlc_i)$ in the left camera coordinate system. Note that the coordinate Zlc_i is the ith gene value in each chromosome.
2. A 3D rotation R_2 is executed on point $Plc_i(Xlc_i, Ylc_i, Zlc_i)$, which provides point $Prc_i(Xrc_i, Yrc_i, Zrc_i)$ in the right camera coordinate system.
3. Prc_i is then transformed into 2D point $P'r_i(X'l_i, Y'l_i)$ in the right image by a perspective projection T_3.
4. A transformation T_4 which takes into account the translation, the residual rotation and the distortion between the two cameras is applied on $P'r_i(X'l_i, Y'l_i)$. Matrix T_4 is calculated by multiplying matrix P which is constituted of all the POI $Pr_{i(1 \leq i \leq 2N)}$ detected on the right image, and the inverse of matrix P', composed of all the points $P'r_i$. Since matrix P' is not a square matrix, we use the pseudo inverse method to calculate its inverse.

Let us consider the following formula:

$$Id = (P'^T P')^{-1} P'^T P' \tag{11}$$

where Id is the identity matrix.

From Eq. 11, we can deduce that T_4 is given by:

$$T_4 = P(P'^T P')^{-1} P'^T \tag{12}$$

Let $Prf_i(Xrf_i, Yrf_i)$ be the transformed point in the right image obtained from $P'r_i(X'l_i, Y'l_i)$ by achieving transformation T_4. Let T be the composition of the all the transformations defined above (namely T_1, R_2, T_3 and T_4). Hence, we have:

$$Prf_i = T(Pl_i) \tag{13}$$

If the depth values in a given chromosome (i.e. the depth of the markers) are accurate, then the two points Prf_i and $Pr_{i(1 \leq i \leq 2N)}$ are equal to the same point for all the couples of matched POI in the left and right images.

The fitness f_j of the jth chromosome is equal to the sum, on the one hand, of all the error distances $d_{i(1 \leq i \leq 2N)}$ between the corresponding points Pr_i and Prf_i, and, on the other hand, of

the maximum error distance calculated in the chromosome. This maximum error value is used in order to penalize some chromosomes which can present a low average value while one or some of the gene values remains very large. The fitness value f_j is then obtained by the following formula:

$$fj = \sum_{i=1}^{2N} (d_i^2)^* \max(d_i) \qquad (14)$$

where N is the marker number.

It should be reminded that the optimization problem at hand is a minimization problem to solve. Thus, the smaller the fitness value, the better the chromosome.

3.2.5 Evolutionary operators

In this section, we describe the different operators that we have defined to allow the population to evolve from a generation to the next.

3.2.6 Selection

A selection by rank [13] is performed in order to select the chromosomes which will take part in the evolutionary operations. First, the chromosomes are sorted from the best to the worst, according to their fitness. Then, a selection probability is assigned to the chromosome with the ith rank as follows:

$$P_i = \frac{2^*(N_p + 1 - i)}{N_P^*(N_P + 1)} \qquad (15)$$

where N_P is the number of individuals of the population.

The advantage of this kind of selection is twofold: the fitness value is not directly taken into account and the selection probability for a chromosome ranked around the average is not insignificant.

3.2.7 Crossover

Two crossover operators are achieved upon the selection in order to diversify the new offspring: a single point crossover and an algebraic crossover.

The single point crossover is a well-known crossover operation (see Fig. 4): a gene is selected randomly in the chromosome. Then, the first child is obtained by taking the beginning part of the first parent, i.e. from its first gene until the selected gene and, the end

Fig. 4 Single point crossover operator
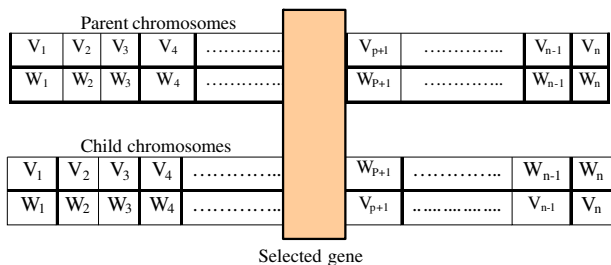
part of the second parent. And the second child is calculated similarly to the first one by taking the end of the first parent and the beginning of the other.

Concerning the algebraic crossover, a linear combination is applied on two chromosomes C1 and C2 in order to obtain two children. The coefficients are calculated with this formula:

$$
\begin{aligned}
\alpha_1 &= \max(r_1, r_2)/(r_1 + r_2) \\
\beta_1 &= 1 - \alpha_1 \\
\alpha_2 &= \beta_1 \\
\beta_2 &= \alpha_1
\end{aligned}
\tag{16}
$$

where the couple (α_1, β_1) (resp. (α_2, β_2)) are the coefficients of C1 and C2 that are utilized in order to calculate the first offspring (resp. the second offspring) and, r_1 and r_2 are the respective ranks of C1 and C2.

The single point crossover allows the production of offspring by combining the genes of two parents in order to improve the future generation while the algebraic crossover provides new gene values by achieving a search between the two hyperplanes defined by the two parents.

The crossover is applied according to a given rate P_c. At the crossover stage we choose between the single point crossover and the algebraic crossover depending on a pre-assigned rate P_{spc} for the single point crossover (respectively P_{ac} for the algebraic one). Note that we have the following formula:

$$
P_{ac} = P_c - P_{spc}
\tag{17}
$$

3.2.8 Mutation

The mutation operation randomly selects one of the genes and replaces its value by a random value chosen in the interval defined for the depth. The mutation is applied depending on the fixed mutation rate P_m. Similarly to the simulated annealing method, we have introduced the temperature as an additional parameter. The temperature decreases at each iteration according to a given schedule such as $T_k = \varphi(T_0, k)$ at the kth iteration. Usually, the initial temperature T_0 is set to one and one takes $\varphi(T_0, k) = a^k T_0$ where a is slightly less than one. At each generation, the mutation rate is multiplied by the temperature. The goal is to decrease the mutation rate as the process progresses. This diminution suppresses the local divergences once the algorithm has converged to the optimal solution.

3.3 Panoramic Reconstruction

3.3.1 2D Surrounding View

The purpose of our 2D surrounding view is to reconstruct an image on which the markers are matched from the leftmost image to the rightmost image. This reconstruction is made just with the translation of the images. It must be mentioned that this is not a real 2D panoramic reconstruction, since no transformation between the images is calculated, despite the different camera angles of view.

The reconstruction process uses at least three consecutive images of the object. The markers are detected in each of the three images. Firstly, the markers on the left image

which are visible on the middle image are matched (see Fig. 7a and b). And secondly, the markers of the right image are similarly matched with the markers of the middle image (see Fig. 7c and b). Then, a 2D transformation is calculated to translate and rotate the left and right images in order to superimpose their POIs with the POIs in the middle image. Finally, each image is vertically cut from the POIs and the corresponding part is retained and juxtaposed for the final image.

3.3.2 3D Panoramic reconstruction

The 3D coordinates of POIs on the different sides of the object are known with the 3D partial reconstruction method, and the 2D surrounding view gives the matching of the POIs between the different points of view. So we can estimate the transformation between the different 3D reconstruction and calculate a 3D panoramic reconstruction of the object.

3D panoramic reconstruction is obtained by a combination of the marker depth values provided by the proposed EA method and the 2D surrounding view as follows:

1. The EA calculates the depth of all the markers which are seen by the different cameras.
2. The 2D surrounding view allows the matching of the points between two consecutive images.
3. The 3D panoramic reconstruction is obtained by superimposing the 3D points corresponding to the matched 2D POIs provided by step 2. This superimposition is performed by aligning the 3D points provided by step 1. A transformation composed of a rotation, a translation and a scale is calculated from these points.

4 Experiments

In this section we describe the experiments which were conducted to objectively assess the performance and the validity of the proposed approach.

4.1 Robustness of the POIs detection and matching processes

The first part of the experiments deals with the detection of POIs and the matching process. Figure 5 shows two series of detection of POIs and the matching results from two different sets of stacking cubes. We added random noise (Fig. 5a) and Gaussian noise (Fig. 5b). Each image shows the result obtained from a pair of images of the same object acquired by two consecutive cameras. Each series shows the result with no noise (a1 and b1), with low noise (a2 and b2) and with high noise (a3 and b3). We can observe that the overall results are correct even with additional noise, which proves the robustness of the method.

4.2 2D Surrounding View

Figure 6 provides the results of the detection of the the POIs using the images of a yellow car. We can note that the highest and the lowest points are well detected on most of the markers. However, there are small location errors concerning the lowest point of the blue marker on the wheel in Fig. 6a and on the lowest point of the right marker in Fig. 6b. Actually, these markers are not well segmented due to the low quality of the images. Figure 6c shows the result of the POIs matching process between the two images illustrated in Fig. 6a and b. It can be observed that all the POIs in the left image are well matched with their respective corresponding POIs in the right image.
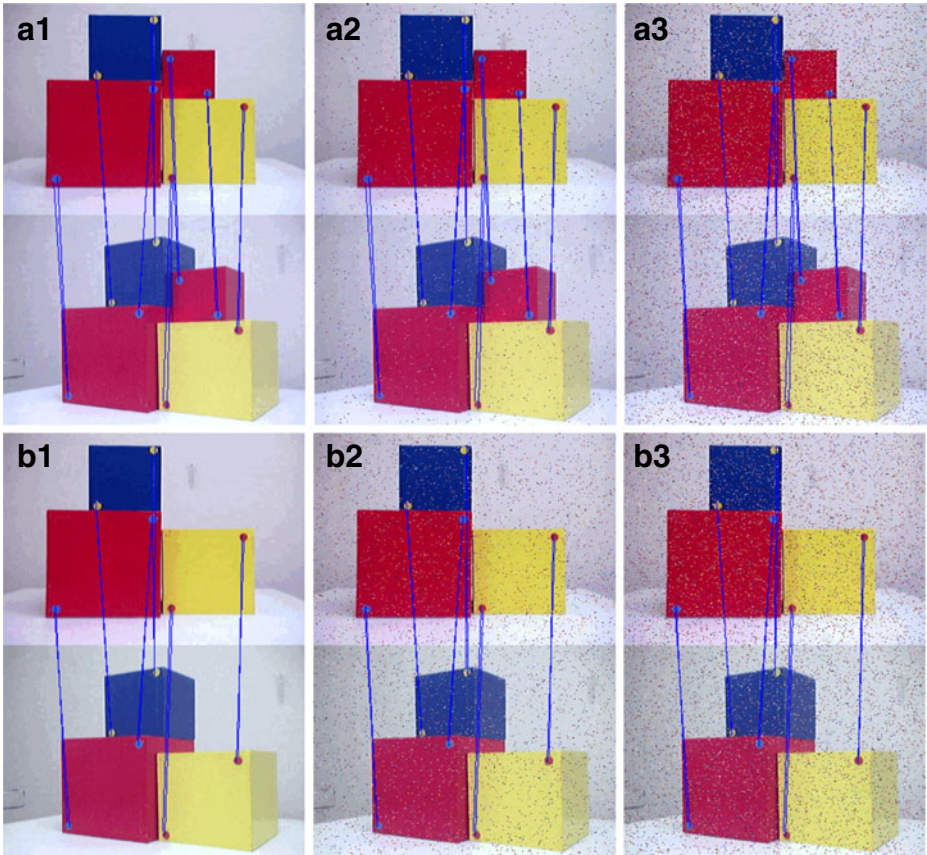
Fig. 5 Illustration of the robustness of the POIs detection and matching method. **a** The object is a stack of four cubes with eight markers. **b** The object is a stack of three cubes with six markers

Figure 7d shows a 2D surrounding view obtained from the images acquired by three consecutive cameras of the USS (see Fig. 7a and c. In this final image, two consecutive images are linked up by superimposing the two corresponding POIs respectively detected in the two images. No transformation is applied on the different parts of the images that compose the final image. Actually, in this view, the matching is only realized for the POIs located on the front of the vehicle. We observe that the result is perfect on this part of the car.

4.3 Parameters and results of our EA

The second part of the experiments was conducted in order to calculate the depth of the markers. Firstly, the EA parameters were determined. And secondly, we validated the estimated depth values.

More than 100 tests were carried out to obtain the most suitable value of each parameter. At each run the optimization process was started with a randomly generated initial population. Since the EA are stochastic, the process was different from one run to another (see Fig. 8). All the runs were carried out until convergence. Table 1 shows the values of the different parameters.
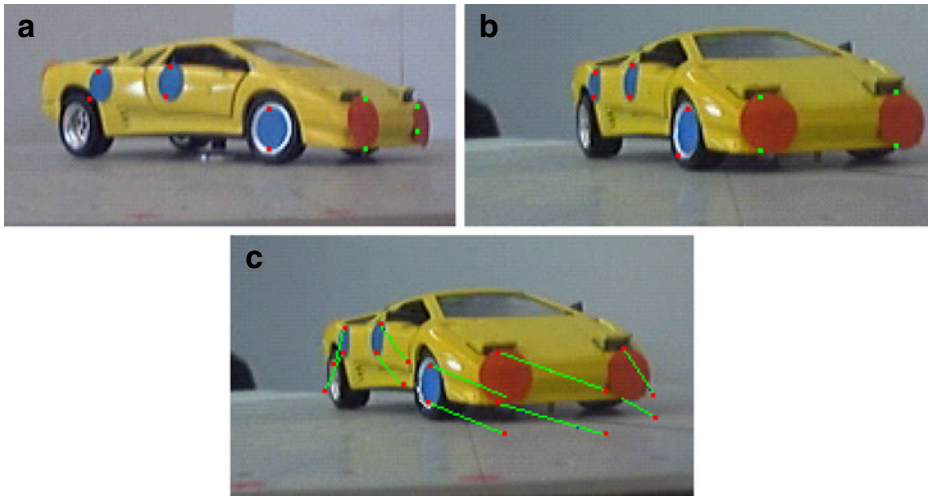
Fig. 6 **a** and **b** Results of POIs detection on two consecutive images. **c** Result of POIs matching - the two corresponding points on the two images are linked by a segment

Figure 8 illustrates an example of the best chromosome fitness curves obtained from three different runs. These curves are representative of most of the curves provided by all the runs. We notice that the process quickly converges to the global minimum. As we can see in Fig. 8a, without any temperature, the algorithm diverges locally all along the process. Conversely, in Fig. 8b and c, we observe that the temperature permits to stabilize the convergence.

Let us notice that in the following, experiment results only concern the yellow car, which is representative of the set of experiments. Ten markers of different colors were placed on the car as follows: two markers on the front and on the back, and three markers on each side. The results shown above give a mean value of about 50 runs.
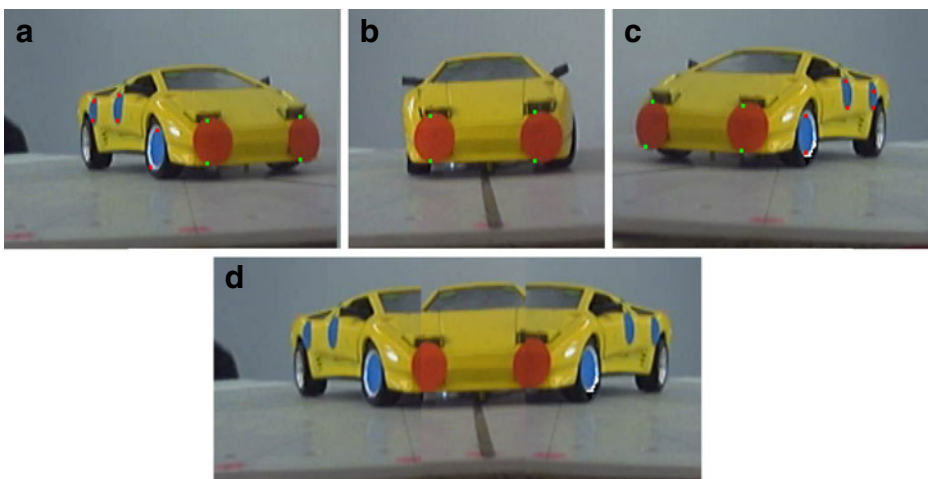
Fig. 7 **a**, **b** and **c** Three different views of the car obtained by three consecutive cameras.(d) The 2D surrounding view from the three images

Fig. 8 Examples of curves of the fitness evolution of the best chromosome vs the generation number: **a** without the temperature parameter, **b** and **c** with the temperature parameter
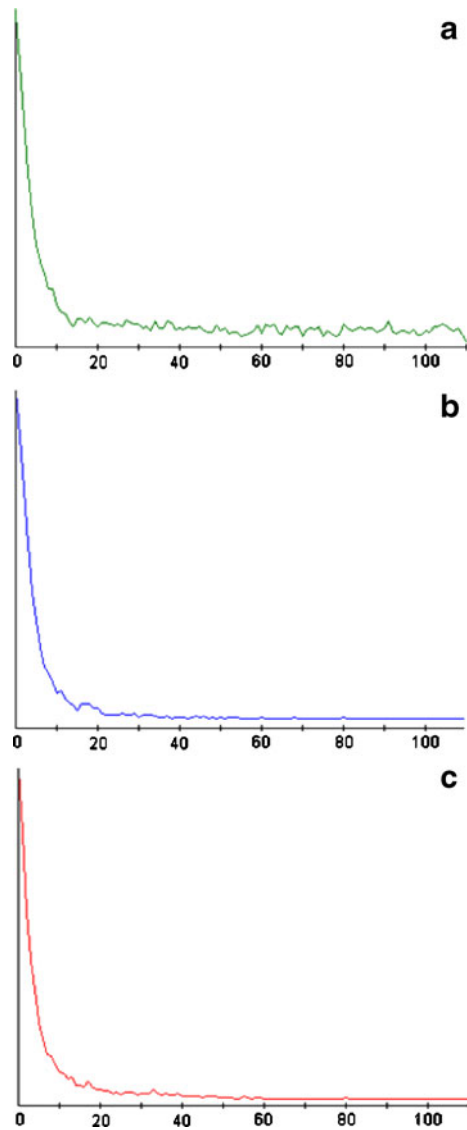
Table 1 EA parameter values

Population fixed-size	100
Number of generations	110
One-point crossover rate (P_{spc})	0.5
Algebraic crossover rate (P_{ac})	0.3
Mutation rate (P_m)	0.3
Temperature decreasing coefficient (a)	0.9
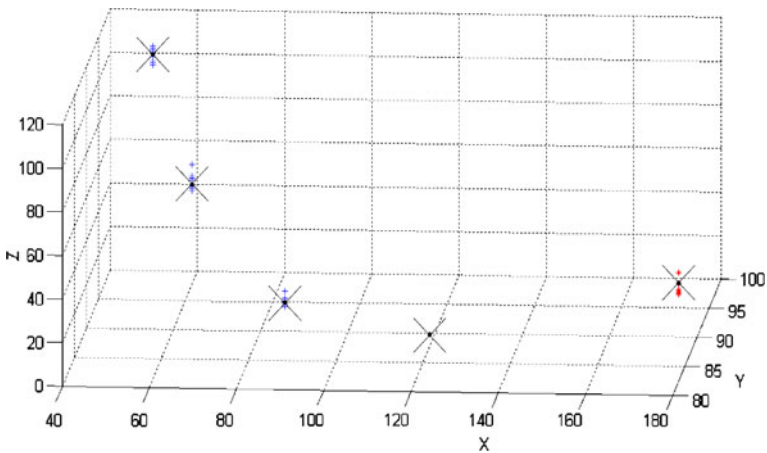
Fig. 9 3D representation of the whole estimated depth results

The origin of the world system is located on the fourth marker, which gives the depth value zero for this marker. On a 2.5 GHz Intel ® Core 2 Duo Processor machine, a run involves approximately 5 s.

Figure 9 illustrates a 3D representation of the obtained depth results from all the 50 runs. Since EA are stochastic, the results from one execution to another may vary. In Fig. 9, the biggest black crosses represent the real values of the marker depths. The little points spread around the real values illustrate the range of all the results. As we can see, the different results are concentrated around the respective real values, which confirms that the EA process is stable and converges well towards the global minimum.

4.4 Comparison with Faugeras-Toscani

In order to validate the depth values obtained with the proposed method, we carried out a comparison with a traditional well known 3D reconstruction method introduced by Faugeras and Toscani [10]. It is important to mention that in our experiments, we chose to compare our results with the Faugeras-Toscani method (FT) because of its robustness and the correctness of its results. Indeed our aim is to prove the validity of the results obtained by our method, but not its effectiveness in terms of computing time.

In the FT framework, the system is calibrated as follows:

1. An image of the calibration pattern is acquired by each camera;
2. The POIs are detected (at least a hundred for a better accuracy) for each camera.
3. The coordinates of the POIs in the pattern are measured in the world system.
4. The calibration matrix between the coordinates in the image and the world system is calculated from Eq. 18 given by Faugeras. Table 2 shows the two respective calibration matrixes for the left camera (Ml) and for the right camera (Mr) in our experiments.

Table 2 The two calibration matrixes

$$[M_l] = \begin{bmatrix} 0.12 & 77.96 & -149.62 & 13548.28 \\ 0.0008 & 27.42 & -52.39 & 4736.24 \\ 0 & 0.46 & -0.88 & 80.40 \end{bmatrix} \qquad [M_r] = \begin{bmatrix} 0.27 & 56.37 & -108.32 & 9818.91 \\ -0.03 & 32.35 & 61.47 & 5559.67 \\ -0.0003 & 0.46 & -088 & 80.71 \end{bmatrix}$$
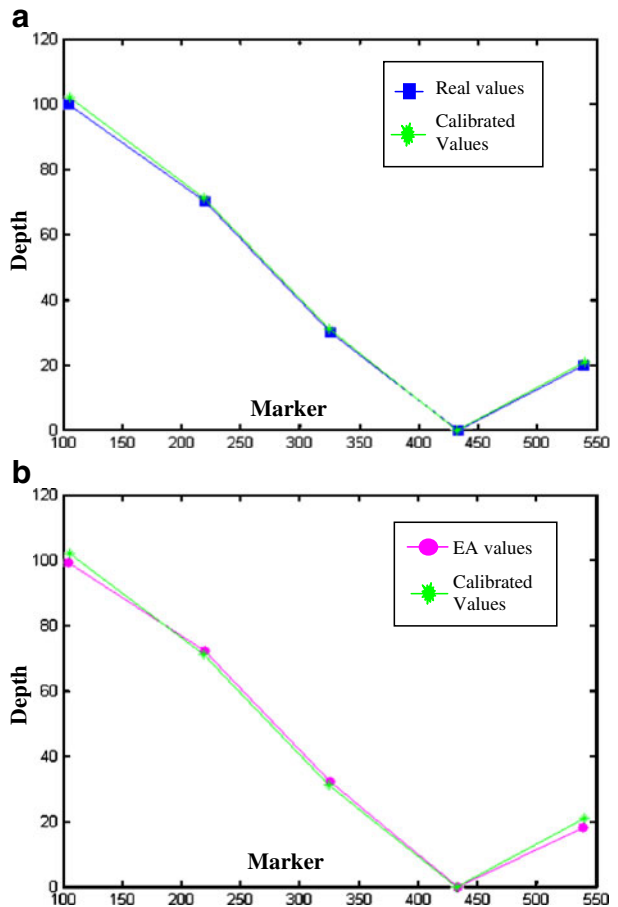
Table 3 Results of the EA and the calibrated methods in comparison with the real values

	Depth marker 1	Depth marker 2	Depth marker 3	Depth marker 4	Depth marker 5
Real values (mm)	100	70	30	0	20
EA method (mm)	99	72	32	0	18
Calibration method (mm)	102	71	31	0	21

5. The extrinsic and intrinsic parameters are extracted.
6. With these parameters the transition matrix from the left to the right camera is calculated.
7. Epipolar geometry provides equations to calculate the depth for all the points.

$$C_{(2N \times 9)} \cdot X_{1(9 \times 1)} + D_{(2N \times 3)} \cdot X_{2(3 \times 1)} = 0 \qquad (18)$$

Fig. 10 a Curves of the real and calibrated depth values. b Curves of the calibrated and EA depth values

Table 4 Comparison between the results of the EA and the calibrated methods

	Execution time (s)
Faugeras-Toscani	1.4
EA method	7.3
Ratio EA/FT	5.21

Where N is the number of POIs detected and X_1 and X_2 are composed as follows:

$$X_1 = \begin{bmatrix} M_1{}^T \\ M_2{}^T \\ m_{34} \end{bmatrix} \tag{19a}$$

$$X_2 = \begin{bmatrix} M_3{}^T \end{bmatrix} \tag{19b}$$

Where M_i are the coefficients of the *ith* line in the matrix M (see Eq. 5).

Table 3 and Fig. 10 illustrate the comparison between the depth values obtained with the proposed EA-based method and the calibrated Faugeras method. We remark that the results provided by the proposed method are globally equivalent to the calibrated method results. Moreover, the accuracy of the respective results obtained by both methods is very satisfactory.

Table 4 shows the execution times of both methods in the case of the yellow car. Faugeras-Toscani method is 5.21 time faster than our EA-based method. But the FT method requires a calibration phase which takes several minutes. Moreover, the camera locations can not be modified after the calibration. Inversely, with our method the calibration step is not necessary which allows the acquisition of images "on a fly". In addition, EA-based methods can easily be parallelized [20, 29], which permits to reduce the computing time.
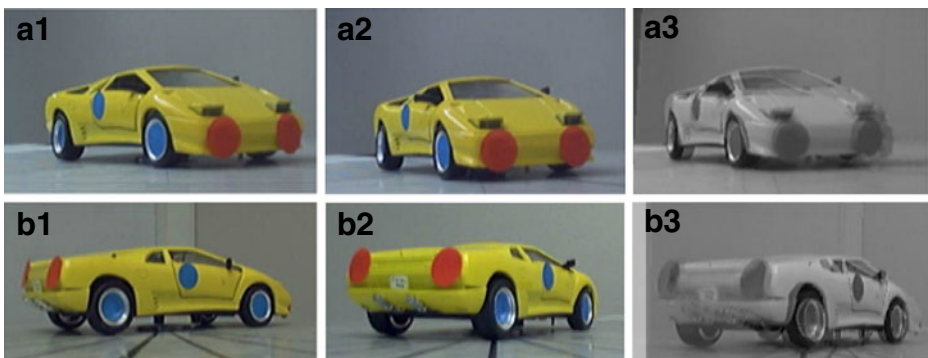
Fig. 11 Two examples of the transformation results. **a1**) and **a2**) or **b1**) and **b2**): two successive images. **a3**) (or **b3**): superimposition of the image in **a2**) (or **b2**) and the transformed image in **a1**) (or in **b1**)

4.5 3D panoramic reconstruction

Figure 11 shows two examples of images obtained by using the calculated transformation T. For each example, the third image ((a3) or (b3)) is the result of the superimposition of the second image ((a2) or (b2)) and the transformation result applied on the first image ((a1) or
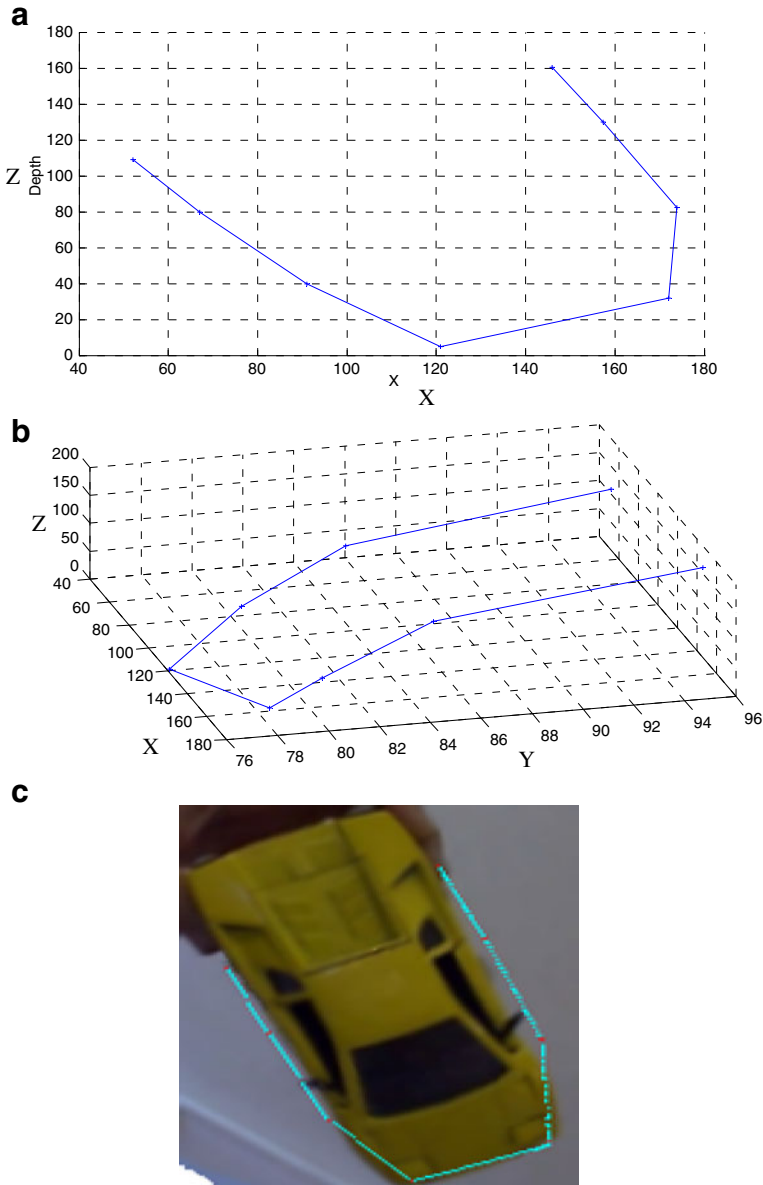
Fig. 12 Two different points of view of the 3D reconstruction of the POIs of the car. **a** a top view of the 3D panoramic reconstruction. **b** a right side point of view. **c** The superimposition of the image of the car acquired from a top point of view and the reconstruction of image (**a**)

(b1)). As seen in both cases, the superimposition results are quite precise on the parts of the car where the markers are entirely visible in the two images. Actually, if a part of the marker is hidden, both points of interest on this marker in the respective two images may not represent exactly the same physical point, which brings about reconstruction mistakes. This can be observed on the right front of the car on Fig. 11c and on the left part of the car on Fig. 11f.

The 3D panoramic reconstruction limits the reconstruction to the points of interest. So in Fig. 12, only the points of interest of the different markers are visible. In order to improve the reading of the result, only the highest point of each marker is illustrated in the figure. The result of the reconstruction is superimposed with a real 2D top view image of the car to have a visual verification of the transformation calculated between the different 3D partial reconstructions. We observe that the result of the 3D panoramic reconstruction is correct.

5 Conclusion and discussion

In this paper we have presented a global image analysis system for reconstructing 2D surrounding views and 3D panoramic images. We use an uncalibrated stereovision system which is composed of five cameras circularly positioned around the object to be analyzed. The proposed method consists of two main stages: firstly, the detection and the matching of some points of interest respectively detected on two images acquired by two successive cameras of the system and secondly, the determination of the interesting point depth values and the transformation matrix between the two images. We have defined evolutionary operators and an original fitness function which are both well-appropriate to solve the problem at hand. Experimental results validate the effectiveness and the correctness of the proposed method in comparison with the results obtained by the well-known calibrated method introduced by Faugeras and Toscani.

It is important to mention that the error on calculating the depth comes from the low resolution cameras used in this study and from the difference between the pin-hole model and the real camera. A way to improve the results is to use a better camera, but we choose to keep the low resolution cameras in order to certify that the results with our method are correct for all the digital cameras.

Our ultimate goal for this work is to allow a complete 3D reconstruction of an object from images taken by a simple camera, and using the location information of the different views, this information being directly provided by a GPS.

Although the study presented in this paper offers very promising results, it still suffers from many drawbacks before leading to applications in static and dynamic domains. Our ongoing work is aimed at carrying out different goals. Firstly, the number of reconstructed points is actually insufficient. Indeed, in the example of the reconstruction of an object shown in the experiments, only five markers are detected. However, we have applied the method to the reconstruction of 20 markers on the stacks of cubes. The accuracy of the results is similar to those shown on the car with longer but acceptable computation times (about 10 s). One of the major problems of EAs is the time of calculation which can be very long in case of a large-sized chromosome. To reconstruct a larger number of points, the image has just to be divided into different parts to be treated independently. Secondly, the number of detected points of interest must be increased. This can be achieved by including a structured light projector into the system in order to create artificial marker whatever the object. And finally, we are studying a new multi-agent evolutionary algorithm in order to improve the computation time.

References

1. Alvarez L, Morales F (1997) Affine morphological multiscale analysis of corners and multiple junctions. Int J Comput Vis 2(25):95–107
2. Armangue X, Salvi J (2003) Overall view regarding fundamental matrix estimation. Image Vis Comput 21:205–220
3. Baker JE (1985) *Adaptive Selection model for genetics algorithms*. Proc. of an international Conference on genetic algorithms and their applications, Hillsdale, Nw Jersey, USA, 101–111
4. Brown MZ, Burschka D, Hager GD (2003) Advances in computational stereo. IEEE Trans Pattern Anal Mach Intell 25(8):993–1008
5. Bunschoten R, Kröse B (2002) 3D scene reconstruction from cylindrical panoramic images. Rob Autom Syst 41:111–118
6. Cantú-Paz E (1998) A survey of parallel genetic algorithms. Calculateurs Parallèles, Réseaux et Systèmes Répartis 10(2):141–171
7. Dai S, Ji Q (2001) A new technique for camera self-calibration. Proceedings of IEEE Conference on Robotics and Automation, Seoul, Korea, pp. 2165–2170
8. Deriche R, Giraudon G (1993) A computational approach for corner and vertex detection. Int J Comput Vis 10(2):101–124
9. Dipanda A, Woo S, Marzani F, Bilbault JM (2003) 3-D shape reconstruction in an active stereo vision system using genetic algorithms. Pattern Recognit 36(9):2143–2159
10. Faugeras O, Toscani G (1987) Camera Calibration for 3D Computer Vision. Proc. International Workshop on Machine Vision and Machine Intelligence, Tokyo, Japan
11. Fiala M, Basu A (2005) Panoramic stereo reconstruction using non-SVP optics. Comput Vis Image Underst 98(3):363–397
12. Fleming PJ, Purshouse PC (2002) Evolutionary algorithms in control systems engineering: a survey. Control Eng Pract 10:1223–1241
13. Goldberg DE (1989) Genetic algorithms in search, optimization & machine learning, Addison-Wesley
14. Grossman E, Santos-Victor J (2000) Uncertainty analysis of 3D reconstruction from uncalibrated views. Image Vis Comput 18:685–696
15. Harris C, Stephens M (1988) A combined corner and edge detector. Proc. of Alvey Vision Conference, pp 147–151
16. Hespanha JP, Dodds Z, Hager GD, Morse AS (1999) What tasks can be performed by an uncalibrated stereo vision system? Int J Comput Vision 35(1):65–85
17. Hill A, Taylor CJ (1992) Model-based image interpretation using genetic algorithms. Image Vis Comput 10(5):295–300
18. Hu M, McMenemy K, Ferguson S, Dobbs G, Yuan B (2008) Epipolar geometry estimation based on evolutionary agents. Pattern Recognit 41(2):575–591
19. Jankó Z, Chetverikov D, Ekárt A (2006) Using a genetic algorithm to register an uncalibrated image pair to a 3D surface model. Eng Appl Artif Intell 19(3):269–276
20. Knoll P, Mirzaei S (2003) Validation of a parallel genetic algorithm for image reconstruction from projections. J Parallel Distrib Comput 63(3):356–359
21. Koljonen J, Alander JT (2006) Genetic algorithm for optimizing fuzzy image pattern matching. Proceedings of the Ninth Scandinavian Conference on Artificial Intelligence (SCAI 2006), Finland, 46–53
22. Koza JR (1995) Survey of genetic algorithms and genetic programming. Proc. of WESCON Conference, 589–594
23. Longuet-Higgins HC (1981) A computer algorithm for reconstructing a scene from two projections. Nature 293(8):133–135
24. Mendonca PRS, Copolla R (1999) A simple technique for self-calibration. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vol. 1, pp. 500-505
25. Menegatti E, Pajdla T (2010) *O*mnidirectional robot vision. Robot Auton Syst 58(6):745–746
26. Meng Y, Zhuang H (2007) Autonomous robot calibration using vision technology. Robot Comput-Integr Manuf 23:436–446
27. Mouaddib E, Sagawa R, Echigo T, Yagi Y (2005) Stereo vision with single camera and multiples mirrors. IEEE International Conference on Robotics and Automation, Barcelona, Spain, pp. 812–817

28. Stefano LD, Marchionni M, Mattocia S, Neri G (2002) A fast area-based stereo matching algorithm. IAPR/CIPRS International Conference on Vision Interface
29. Wang ZG, Wong YS, Rahman M (2005) Development of a parallel optimization method based on genetic simulated annealing algorithm. Parallel Comput 31(8–9):839–857
30. Yang F, Paindavoine M, Abdi H, Arnoult D (2006) Fast image mosaicing for panoramic face recognition. J multimed 1(2):14–20
31. Yin X, Xie M (2003) Estimation of the fundamental matrix from uncalibrated stereo hand images for 3D hand gesture recognition. Pattern Recognit 36:567–584
32. Zhang J, Ge Y, Ong SH, Chui CK, Teoh SH, Yan CH (2008) Rapid surface registration of 3D volumes using a neural network approach. Image Vis Comput 23:201–210
33. Zhengzhou Z (2000) A flexible new technique for camera calibration. IEEE transactions on pattern analysis and machine intelligence, pp. 1330–1334

Alain Koch received his Master degree in computer science from the University of Burgundy, in France, in 2006. He is currently student at the same University, obtaining his Ph.D. degree in computer vision and image processing. His research interests include 3D reconstruction with uncalibrated system and evolutionary algorithm.

Albert Dipanda received his Ph.D. degree in computer science in 1990 from the University of Burgundy. France. He is currently Professor in the Engineering School of Materials, Computer Sciences and Electronics and Deputy Director of the LE2I laboratoy (Laboratoire d'Electronique, Informatique et Image). His research interests include motion estimation, MRF modeling in Image Processing, 3-D reconstruction and Genetic Algorithm applications in image processing.

Claire Bourgeois République received her Ph.D. in Computer Sciences from the University of Burgundy, France in 2004. She is an assistant professor at the University of Burgundy since 2005. Her research interest domain was the Artificial Life (genetic algorithms), 3-D reconstruction and Genetic Algorithm applications in image processing. She joint the Image Processing group of the LE2I (Laboratoire d'Electronique, Informatique et Image: Laboratory of Electronics, Computer Sciences and Image) in 2004.