Coral Reef Optimization for Intensity-based Medical Image Registration

Enrique Bermejo*, Manuel Chica†, Sancho Salcedo Sanz‡ and Oscar Cordón*
*Dept. of Computer Science and Artificial Intelligence, University of Granada, Spain
{enrique.bermejo, ocordon}@decsai.ugr.es
†School of Electrical Engineering and Computing, University of Newcastle, Australia
manuel.chica.serrano@gmail.com
‡Dept. of Signal Theory And Communications, University of Alcalá, Spain
sancho.salcedo@uah.es

***Abstract*—Image registration (IR) is an extended and important problem in computer vision. It involves the transformation of different sets of image data having a shared content into a common coordinate system. Specifically, we will deal with the 3D intensity-based medical IR problem where the intensity distribution of the images is considered, one of the most complex and time consuming variants. The limitations of traditional IR methods have boomed the application of evolutionary and metaheuristic-based approaches to solve the problem, aiming to improve the performance of existing methods both in terms of accuracy and efficiency. In this contribution, we consider the use of a recently proposed bio-inspired meta-heuristic: the Coral Reef Optimization Algorithm (CRO). This novel algorithm simulates the natural phenomena underlying a coral reef, where different corals grow, reproduce and fight with other corals for space in the colony. CRO has recently obtained promising results in different real-world applications and we think its operation mode can properly cope with the 3D intensity-based medical IR problem. We adapt the algorithm to the real-coding problem nature and run an experimental setup tackling sixteen real-world problem instances. The new proposal is benchmarked with recent, state-of-the-art IR techniques. The results show that the CRO-based overcomes the state-of-the-art results in terms of its robustness and time efficiency.**

I. Introduction

In computer graphics, a large number of applications involve the alignment of multiple images having a shared content. This task is known as Image Registration (IR) [1] and it is focused on finding the optimal geometric transformation that overlaps the common part of the images. IR is an essential task in medical image analysis, as it allows the integration of visual information obtained under different conditions [2]. Among the clinical applications of this task, we find: combining different images of the same patient, detecting changes over time (before/after treatment), aligning temporal image sequences to compensate for motion between scans, image guidance during interventions, and image alignment from multiple patients in cohort studies.

Usually, IR methods consider a spatial transformation to align images by mapping their overlapping areas. The alignment procedure either makes use of the entire image (intensity-based approaches) or it is based on salient or distinctive parts of the images (features). Feature-based approaches [1], [2] are faster as they notably reduce the complexity of the problem by only using relevant information extracted from the images. However, they are highly dependent on the feature extraction process, a stage that may lead to inevitable errors if the features are not able to provide representative information of the images. At the expense of increasing the computational requirements, intensity-based (or voxel-based) approaches cannot only overcome these problems, but they also provide the most accurate and robust outcomes [3]. In medical applications, intensity-based is the most common approach as these applications require high levels of precision and consistency. Thus, the alignment is guided by the distribution of intensity values (gray levels) of the whole images.

In most of the IR approaches, the problem of estimating the spatial transformation between images is treated as an iterative optimization procedure that explores the space of possible transformations. The quality of a solution is defined by the degree of resemblance between images after the transformation, which is measured by a similarity metric [3].

Traditional IR methods, such as the Iterative Closest Point algorithm [4] and different gradient-based approaches [5], [6] are highly influenced by different factors: noise, discretization, and misalignment, among others. Methods based on evolutionary algorithms and other metaheuristics (MHs) [7] are able to overcome some of these drawbacks. Thus, MH approaches are often considered to tackle medical IR problems due to their robust performance.

In [8], the authors included a thorough comparison of several methods applied to different IR problems. They proposed a novel evolutionary approach based on scatter search (SS), which is currently the state of the art. Later, in [9] we proposed an alternative approach based on the bacterial foraging optimization algorithm (BFOA) to compete with outstanding IR methods. The results showed that our adaptation of BFOA was competitive but its performance was inconsistent and unable to outperform SS.

Recently, a novel bio-inspired evolutionary-type MH called coral reef optimization algorithm (CRO) was proposed in [10]. CRO is an evolutionary algorithm based on the artificial simulation of the natural phenomena taking place in the formation and reproduction of coral reefs. During their life, corals undergo different phases, such as reproduction, larval settlement, or fight for a space in the reef where they can survive. The proposed CRO approach emulates these processes favoring a powerful trade-off between diversity and specificity,

978-1-5090-4601-0/17/$31.00

which makes it suitable for tackling complex optimization problems.

Since CRO has been successfully applied to different real-world problems by demonstrating a robust performance, we consider it can also be applied to complex IR problems and behave adequately. Thus, in the present work, we design a novel IR method based on CRO in which both the objective function and the coding scheme were adapted to our specific medical IR problem.

In order to validate our proposal, we develop an exhaustive experimental setup comparing CRO against the state-of-the-art IR methods in sixteen real-world scenarios. Each problem instance is generated using pairs of images extracted from the well-known BrainWeb database at McGill University [11]. Different statistical tests are applied to evaluate the significance of the results.

The paper is structured as follows: In Section II, the IR problem tackled in this paper is introduced, detailing and analysing the most relevant IR existing methods and specifically focusing on the 3D intensity-based medical IR problem. Section III describes the basics of the CRO algorithm and its adaptation to solve our IR problem. Section IV introduces the test scenarios, the experimental comparison, and the analysis of their results. Finally, Section V presents some conclusions and remarks on the research carried out.

II. Image Registration

A. Problem statement

In medical IR problems, we are usually provided with two images, a reference image, also called *model* (I_M), and the image that is transformed to reach the model geometry, known as *scene* (I_S). The objective of the registration process is to find a geometric transformation T making the model I_M and the transformed scene $T(I_S)$ be as similar as possible. The degree of resemblance between the considered images is measured using a similarity metric. Hence, the IR problem can be formulated as a maximization problem over the space of transformations:

$$\underset{T \in \text{Transformations}}{\operatorname{argmax}} \quad \text{Similarity}\,(I_M, T(I_S)).$$

Any IR method involves three main components: the transformation model, the similarity metric, and the optimization procedure. Figure 1 shows the flow chart of a typical iterative optimization process. First, the optimizer estimates a candidate transformation, which is applied to the scene image. Then, the similarity metric measures the quality of the alignment. This result is then sent back to the optimizer, creating a loop that ends up when a particular stopping criteria is reached (e.g., when a suitable solution has been found or the algorithm performs a certain number of iterations). We explain these three components with more details in the following paragraphs:

1) Transformation model: The first component to consider is the transformation model, which determines what kind of geometrical transformation is used to align the images. The appropriate transformation model entirely depends on both the specific application and the nature of the images involved.

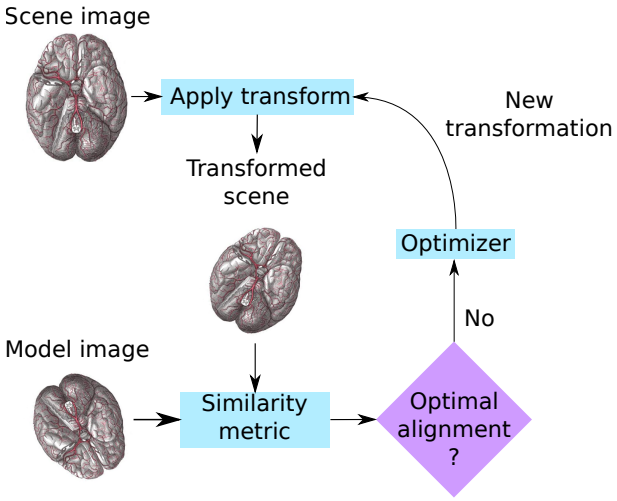


Figure 1. The interactions among the three components (blue boxes) of a registration technique.

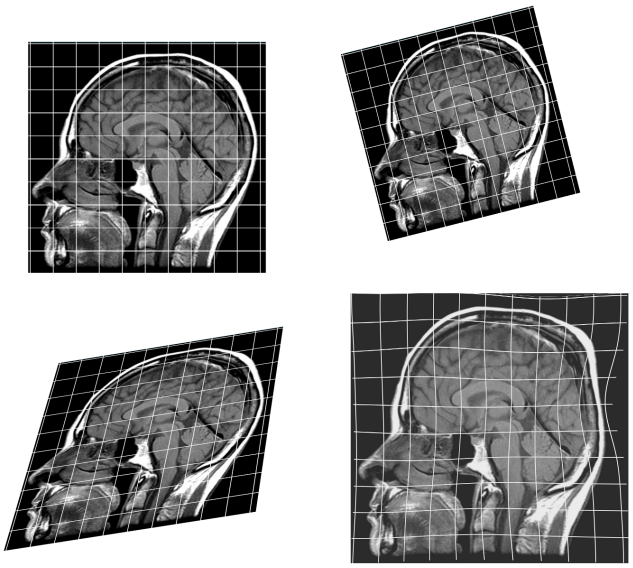


Figure 2. Three images obtained from the same initial scene (top left) by applying different transformations: similarity (top right), affine (bottom left), and B-spline (bottom right).

In certain contexts, such as remote sensing, a simple model (e.g., a translation transformation) can be an adequate choice, while other applications involving motion estimation require deformable models. Anyhow, the specific kind of transformation has to be carefully chosen, not only in terms of its number of parameters determining the computational effort to estimate them, but also considering the resulting transformation effects. For instance, in medical applications, bent bones or unrealistic tissue growth are undesired consequences of the application of too flexible transformation models. Figure 2 illustrates the effects of applying different transformations to a specific scene.

Usually, rigid transformations are considered in most of the medical IR applications which involve anatomical structures like bones or the brain. A rigid transformation allows translation and rotation movements, but there are applications that

also involve scaling. In these cases, a similarity transformation is used to cope with three movements: translation, rotation, and scaling. More complex scenarios, particularly those involving CT scans, may require affine transformations, due to the presence of shearing (i.e., non-uniform scaling in some directions). In applications where soft tissue is involved, a significant deformation can be expected. Therefore, non-rigid or deformable transformations must be used.

2) Similarity metric: The similarity metric is an essential component of any IR method [12] as its performance depends on the precise estimation of the alignment between images. The quality of a transformation T is obtained by computing the similarity metric over the model I_M and the transformed scene $T(I_S)$ images, noted as $F(I_M, T(I_S))$. The technique used to evaluate the resemblance between both images depends on the considered registration approach. For instance, in feature-based methods, the similarity metric usually evaluates the quality of the alignment by measuring the distance between corresponding features [13]. Thus, each point x_i in the model is paired with the closest point in the transformed scene, and the mean square error (MSE) is calculated:

$$\text{MSE} \;=\; \frac{1}{r}\sum_{i=1}^{r} \|x_i - c(T(x_i))\|^2,$$

where r is the number of points and c is the correspondence criterion.

Alternatively, in intensity-based approaches like the one tackled in the current contribution, the similarity metric considers the resemblance of the intensity values in the two images. The suitability of a similarity metric depends on the relationship between the intensity distributions of the images, which depends in turn on the subject of the images and their acquisition technique. Multiple metrics are available, though common choices are sum of squared differences, normalized correlation (NC), and mutual information [14]. In particular, NC is used to assess the similarity when the relationship between images is linear. The linear NC is defined as:

$$\text{NC}(I_A, I_B) = \frac{\sum_{x\in\Omega} \left(I_A(x) - \overline{I_A} \right) \left(I_B(x) - \overline{I_B} \right)}{\sum_{x\in\Omega} \left(I_A(x) - \overline{I_A} \right) \sum_{x\in\Omega} \left(I_B(x) - \overline{I_B} \right)},$$

where $\overline{I_A}, \overline{I_B}$ are the average intensity values of images A and B, and Ω is the common part of their domain. Nonetheless, other scenarios can be strongly non-linear, such as the case when the images have been acquired using different sensors and their intensity distribution is different. These scenarios are known as multimodal registration, and metrics based on information theory are the most suitable option, such as the mutual information measure. This measure is defined as:

$$\text{MI}(I_A, I_B) = \sum_{a\in I_A}\sum_{b\in I_B} p_{AB}(a,b) \log \frac{p_{AB}(a,b)}{p_A(a)\; p_B(b)},$$

where p_{AB} is the joint probability and p_A, p_B are the marginal discrete probabilities of the intensity values of the images.

3) Optimization procedure: The optimization procedure is a crucial part of any IR method. It consists of an iterative procedure responsible for exploring the space of geometrical transformation, guided by the similarity metric. Two strategies are available. In *parameter-based* approaches, the search is performed directly in the space of the transformation parameters, as it considers the registration as a continuous optimization problem. Classic numerical optimization algorithms such as gradient descent, conjugated gradient descent, Newton's and quasi-Newton methods, Powell's method and discrete optimization [15], [16] are among the most common choices for the optimization component. The use of approaches based on evolutionary computation and other MHs to deal with the drawbacks presented by classical search methods [17]–[19] is very common. Alternatively, in *matching-based* approaches the search can be focused on the space of feature correspondences. The correspondence can be done either by considering matching features (*feature-based* methods) or by comparing areas of the image (*intensity-based* methods). The parameters of the transformation can be extracted from the estimated matching using least-squares estimation or other model fitting techniques [20]. A classical matching-based algorithm is the Iterative Closest Point [21].

B. Current best performing algorithms for medical IR

The two best performing methods reviewed here are MH-based approaches, which will be used for comparison purposes during the experimentation. In addition to the methods described below, three different alternative approaches are described in [9] considering both classical gradient-based algorithms and different MHs. They will be also considered for the comparison in the experimental section.

a) Bermejo et al.'s h-BFOA: BFOA is a swarm intelligence algorithm inspired by the foraging strategy of *E. Coli*. The objective of this strategy is to search for nutrients in the environment, maximizing the amount of energy obtained in the process. Beides, an individual bacterium can communicate with others indicating food sources or possible hazards. h-BFOA [9], [22] is a novel design of the canonical algorithm proposed by [23], adapted to intensity-based IR. The hybridization incorporates a local search (LS) strategy, applied to the best bacterium of each generation. We considered a Crossover Based Local Search, known as XLS [24], which is based on auto-adaptive crossover operators. Thus, when a parent solution (best bacterium) is given, XLS selects L random bacteria in the population and generates new ones in the parent neighborhood.

b) Valsecchi et al.'s SS: SS$^+$ [8] is based on a variant of the original SS design [25] in which the reference set is divided in two tiers, containing the highest quality and diverse solutions, respectively. This variant was designed specifically for IR, selecting the best performing methods from the SS template. SS$^+$ integrates a diversification generation method, which is an approach based on frequency memory [26], to ensure that the search space is explored in a uniform manner. In addition, the solution combination method is the BLX-α crossover, while the improvement method is based on the PMX-α [27] operator, a 'parent-centric' version of the latter. The reference set update method ensures the highest-quality solutions are maintained in the quality reference set while the most diverse solutions remain in the diversity reference set. Last, SS$^+$ includes a duplication control method, which

prevents identical copies of the same solution from dominating the reference set.

Both approaches consider multiple resolutions combined with a restart mechanism. These optimizer components aim to speed up the optimization process and mitigate a premature convergence of the algorithm (see Section III for a detailed explanation).

III. A Coral Reefs Optimization Algorithm for intensity-based medical image registration

The CRO algorithm [10], [28] is an evolutionary-type algorithm based on the behavior of the processes occurring in a coral reef. Let $\mathcal{R}$ be the reef represented by an $R_1 \times R_2$ grid, where each position (i, j) of $\mathcal{R}$ is able to allocate a coral or a colony of corals, $\mathcal{C}_{i,j}$, standing for solutions to the current optimization problem at hand. Therefore, a coral encodes a candidate solution to the optimization problem. In our specific IR approach, this solution is encoded as a real-valued vector that represents the parameters of a transformation model. The quality of a solution is evaluated using a fitness function. Our approach has been designed so that any similarity metric can be used directly as fitness function. Thereby, our proposal is able to accommodate different transformation models and similarity metrics providing a great flexibility when designing the experimental setup. Specifically, for this contribution we have considered a similarity transformation, encoded using seven parameters: rotation versor $(\theta_x, \theta_y, \theta_z)$, translation vector (t_x, t_y, t_z), and uniform scaling factor s. Besides, we have chosen the NC metric (defined in Section II) as fitness function to guide the optimization process.

The CRO algorithm first initializes some random positions of $\mathcal{R}$ with random corals, and leaves some other positions empty. These holes in the reef are available to host new corals that will be able to freely settle and grow in later phases of the algorithm. The rate between free/occupied positions in $\mathcal{R}$ at the beginning of the algorithm is a parameter of the CRO algorithm, denoted as ρ_0 with $0 < \rho_0 < 1$.

The second phase simulates the processes of reproduction and reef formation. The different reproduction mechanisms available in nature are recreated by sequentially applying different operators:

1. **External sexual reproduction or Broadcast Spawning.** Broadcast spawning consists of the following steps at each iteration k of the algorithm:

   1.a. A random fraction of the existing corals is selected uniformly, turning these corals into broadcast spawners. The fraction of broadcast spawners with respect to the overall amount of existing corals in the reef will be denoted as F_b.

   1.b. Several coral larvae are formed. To generate each new larva, two broadcast spawners are selected and a crossover operator or any other exploration strategy is applied. Note that once two corals have been selected to be the parents of a larva, they are not chosen anymore at iteration k for reproduction purposes. Corals' selection can be done randomly, uniformly, or using any fitness proportionate selection approach (e.g. roulette wheel).

2. **Internal sexual reproduction or Brooding.** Hermaphrodite corals mainly reproduce by brooding. This reproduction is modelled by means of any kind of mutation mechanism and takes place on a fraction of corals of $1 - F_b$. A percentage P_i of the coral is mutated.

3. **Larvae setting.** Once the larvae are formed either through external or internal reproduction, they will try to set and grow in the reef. Each larva will randomly try to set in a position (i, j) of the reef and, if the location is free, it will set. If the location is already occupied, the new larva will set only if its health function (fitness) is better than that of the existing coral. Moreover, the CRO algorithm defines a parameter η that determines the maximum number of tries a larva can attempt at each iteration k.

4. **Asexual reproduction.** Corals reproduce asexually by budding or fragmentation. The CRO models this mechanism in the following way: the whole set of corals in the reef are sorted according to their level of health value (given by $f(\mathcal{C}_{ij})$). Then, a small fraction (denoted as F_a) of the available corals are duplicated and mutated (with probability P_a) to provide variability, and try to settle in a different part of the reef as in Step 3.

5. **Depredation.** Corals may die during the reef's formation. Therefore, at the end of each reproduction iteration k, a small number of corals in the reef can be depredated, thus liberating space in the reef for next coral generation (iteration $k + 1$). The depredation operator is applied with a very small probability (P_d) to a fraction (F_d) of the corals in the reef with worse health.

Algorithm 1 illustrates the flowchart diagram of the CRO algorithm, with the different CRO phases (reef initialization and reef formation), along with all the operators described above.

Furthermore, following the design of h-BFOA and SS$^+$, the scheme of CRO also integrates two components specific to the IR problem we address:

1. **Multi-resolution strategy** In order to reduce the computational cost of the registration, the input images are pre-processed before the registration, applying both down-sampling and Gaussian smoothing to create two image representations, called pyramids. In the first resolution, the optimizer selects a small, low-detail pyramid of the input images, which allows the algorithm to obtain an approximation of the desired transformation. The second resolution is meant to be a refinement phase, improving the quality of the best transformations found during the first resolution.
2. **Restart mechanism** The main objective of the restart mechanism is to ensure the process can recover from

Require: Valid values for the parameters controlling the CRO algorithm
Ensure: A single feasible individual with optimal value of its *fitness*
1: Initialize the algorithm
2: **for** each iteration of the simulation **do**
3: Update values of influential variables: predation probability, etc.
4: Sexual reproduction processes (broadcast spawning and brooding)
5: Settlement of new corals
6: Predation process
7: Evaluate the new population in the coral reef
8: **end for**
9: Return the best individual (final solution) from the reef
Algorithm 1: Pseudo-code for the CRO algorithm

stagnation and find a good final solution. This mechanism is applied only during the first resolution, where the low-detail pyramids are considered. Thus, the whole sequence of the CRO optimization procedure (Algorithm 1) is performed a fixed number of times, restarting the reef population at the end of each run and storing the best solution found. This solution is carried out to the second resolution, where the refinement is performed. As the first resolution is using a reduced version of the data, this mechanism is computationally cheap.

Algorithm 2 shows the integration of the general structure of CRO with the specific components of IR.

1: Use first level of image pyramids
2: **for** r $\leftarrow$ 1 **to** *Number of Restarts* **do**
3: Run CRO optimization procedure (Algorithm 1)
4: Store best individual (preliminar solution) from the reef
5: **end for**
6: Use second level of image pyramids
7: Run CRO optimization procedure using the best stored individuals during the reef initialization
8: Return the best individual (final solution) from the reef
Algorithm 2: Pseudo-code of the proposed CRO optimization approach for IR.

IV. Experimental study

A. Experimental setup

We design an experimental study by considering the best performing state-of-the-art IR algorithms of [9]: Valsecchi et al.'s SS$^+$ [8], and Bermejo et al.'s h-BFOA [29]; including our novel proposal based on CRO. Table I shows the parameters used in the considered algorithms.

The images used in this experiment are the same used in [9], which were obtained from the BrainWeb database at McGill University [11], consisting in *simulated* brain magnetic resonance images (MRI). The BrainWeb repository is well-known by the IR research community [30]. In order to consider

Table I. Parameter configuration used in CRO, h-BFOA, and SS$^+$.

CRO		h-BFOA		SS$^+$	
Param.	**Value**	**Param.**	**Value**	**Param.**	**Value**
$\mathcal{R}$	150	Swarm size	30	P size	12
ρ_0	0.6	Chemotactic steps	8	RefSet size	8
F_a	0.015	Swim steps	10	Blend factor α	0.3
P_a	0.6	Reproductive steps	18	PMX α	0.5
F_b	0.9	Elimination steps	12	PMX iterations	10
F_d	0.04	Prob. of Elimination	0.125		
P_d	0.1	λ	0.005		
k	3				

Table II. Characteristics of the brain MRI images used in our experimental study, including Gaussian noise level and presence of lesion.

Image	Lesion	Noise
I_1	No	None
I_2	No	1%
I_3	Yes	1%
I_4	Yes	5%

different degrees of complexity, some images include noise and multiple sclerosis lesions, as detailed in Table II. All the images have the same size ($60 \times 181 \times 217$ voxels), as shown in Figure 3.

In order to create different IR problem scenarios, each image was transformed using one of the four similarity transformations (involving rotation, translation, and uniform scaling). The parameters of these transformations are shown in Table III. Note that these parameter values of the transformations produce strong changes in the object location, generating complex IR scenarios. In total, sixteen IR problem instances were created by pairing images with different transformations. From lower to higher complexity, the scenarios are I_1 versus $T_i(I_2)$, I_1 versus $T_i(I_3)$, I_1 versus $T_i(I_4)$ and I_2 versus $T_i(I_4)$, for $i = 1, 2, 3, 4$.

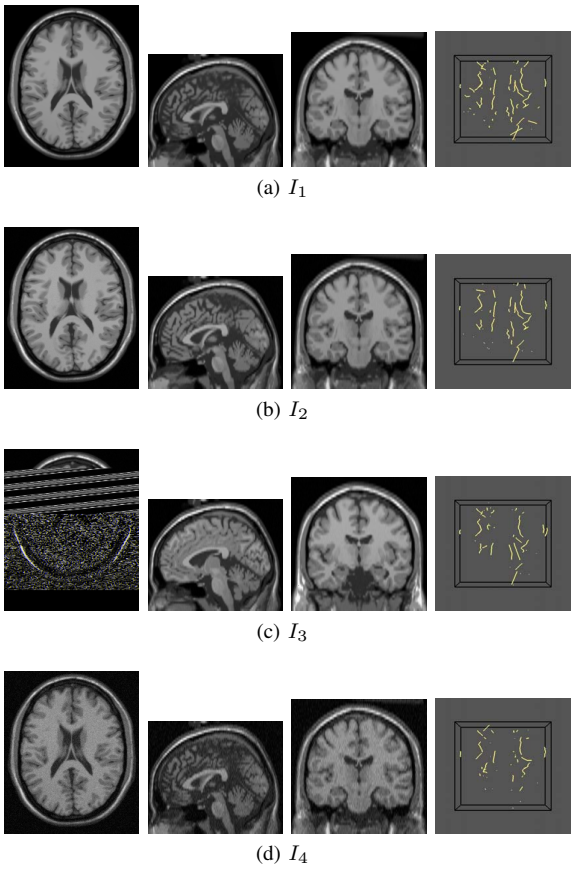
All the algorithms consider similarity transformation as the transformation model, with $[-30, 30]$ as the parameters' range for the translation component, and $[0.75, 1.25]$ for the scaling factor. We did not apply any restriction to the rotation axis. Therefore, the actual range is $[-1, 1]$ for each component of the versor.

IR algorithms may differ in their optimization procedure, and particularly, in the similarity metric, which measures the quality of a solution. Therefore, we considered a common measure that allowed us to evaluate the results among different IR approaches. Thus, the final quality of a solution is obtained by measuring the MSE between the corresponding crest-line points. This set of points with relevant curvature information was extracted using the method described in [31].

In [9], we augmented the common experimental analysis by comparing both the mean and median MSE results obtained

Table III. Parameters of the transformations used in our experiments: rotation angle (λ), rotation axis (a_x, a_y, a_z), translation vector (t_x, t_y, t_z) and uniform scaling factor s.

	λ	a_x	a_y	a_z	t_x	t_y	t_z	s
T_1	115	-0.863	0.259	0.431	-26	15.5	-4.6	1
T_2	168	0.676	-0.290	0.676	6	5.5	-4.6	0.8
T_3	235	-0.303	-0.808	0.505	16	-5.5	-4.6	1
T_4	276.9	-0.872	0.436	-0.218	-12	5.5	-24.6	1.2



(a) I_1

(b) I_2

(c) I_3

(d) I_4

Figure 3. From left to right, axial, sagittal and coronal views, along with the corresponding crest line points used as features of the four considered MRI brain images.

by each considered algorithm. The main reason behind this decision was to provide an unbiased analysis. As optimization algorithms may fail to converge to an acceptable solution in some occasions, it is common that some outliers distort the 'average' meaning of the mean value used for comparison. In this work, we maintain this decision.

For each registration scenario we performed 30 independent runs of each algorithm. As mentioned in Section II-B, the registration was performed in two resolutions. During the first one, the input images are down-sampled and smoothed, and the similarity metric was computed over 3,000 random spatial samples. On the second resolution, no preprocessing was applied, and the number of spatial samples was increased up to 10,000. Note that our goal is to obtain better solutions in less time. Consequently, we limited the running time to 180 seconds, as the fast convergence rate of CRO may satisfy this goal. For the execution of the experiments we used a computer with an Intel Core i7-4790 3.6 GHz processor and 16 GB RAM.

B. Analysis of the results

First, we include for comparison a summary of the results obtained in our previous work [9], where h-BFOA was compared against SS$^+$, a genetic algorithm (r-GA$^+$) proposed in [32], and Klein et al.'s adaptive stochastic gradient descent (ASGD) [6]. h-BFOA showed a competitive performance regarding median MSE results, but had an inconsistent behavior when considering mean MSE values (Table IV). On the contrary, SS$^+$ registered a robust behavior in terms of both median and mean MSE values and confirmed its dominance as state-of-the-art IR method.

Table IV. SUMMARIZED RESULTS OF THE EXPERIMENT FROM [9]. THE TABLE REPORTS THE MEDIAN, MEAN, STANDARD DEVIATION, MEDIAN RANK AND MEAN RANK OF MSE VALUE. THE RESULTS ARE AVERAGED AMONG THE SIXTEEN CONSIDERED INSTANCES.

	MSE			Ranking	
Algorithm	median	mean	stdev	median	mean
ASGD	51518.1	51494.6	566.0	4.00	4.00
h-BFOA	**40.3**	88.4	242.6	**1.50**	2.50
r-GA$^+$	42.8	42.8	0.5	2.37	1.88
SS$^+$	42.7	**42.7**	0.0	2.12	**1.62**

Next, we include in Table V the results of the experiments developed in the current contribution, considering our novel proposal, CRO. We computed median, mean, and standard deviation of the MSE values for each IR scenario. Last, we provide an overall view of the results by including the average ranking of the algorithms regarding all the scenarios. To that end, we apply different post-hoc statistical analysis, which will allow us to detect significant differences among the algorithms. In particular, we consider Friedman's nonparametric test [33], Bonferroni-Dunn's test [34] and Holm's test [35], analyzing the statistical results for both median and mean MSE values. Table VI shows the average ranking and the statistical results regarding the median MSE values, while Table VII refers to the ranking and statistics considering mean MSE values.

CRO had excellent results according to the median and mean MSE values, obtaining the best averaged rank in both cases with 1.68 and 1.06 values, respectively. Considering median MSE, the algorithm outperforms both SS$^+$ and h-BFOA in 7 out of the 16 scenarios. In terms of mean MSE, our proposal obtained the best results in all but one scenario.

Overall, SS$^+$ obtained a good and robust performance in most of the scenarios, behaving consistently regarding median values, and obtaining the best mean results in one of the 16 instances. Even though its performances are good, SS$^+$ obtained the worst average rank (with a 2.5 value) considering median MSE, and second average rank (2.37) in terms of mean MSE.

The performance of h-BFOA greatly varies depending on the scenario, since the algorithm is not able to converge properly in many runs as a consequence of the run time reduction. Considering the median MSE, the algorithm showed a consistent performance, competitive when comparing against the other algorithms and ranking second (1.81) concerning the average rank. But in terms of mean MSE, the algorithm failed to converge steadily, obtaining considerably poor results in 6 of the 16 scenarios. This leads h-BFOA to obtain the worst rank with a 2.56 value.

It is noticeable that reducing the running time from 20 to 3 minutes affected negatively both SS$^+$ and h-BFOA algorithms. SS$^+$ loses consistency, increasing the standard deviation values, while mean MSE results obtained by h-BFOA are skewed due to the increase of outliers. On the contrary, CRO proved to be an extremely fast and robust optimization procedure when tackling this IR problem variant, obtaining better results in 3

Table V. DETAILED RESULTS (MEDIAN, MEAN AND STANDARD DEVIATION) FOR THE MSE VALUES.

		MSE		
	Algorithm	median	mean	stdev
1	h-BFOA	**36.7**	48.7	51.1
	SS$^+$	37.2	38.8	5.0
	CRO	**36.7**	**38.7**	10.8
2	h-BFOA	**36.9**	45.3	30.9
	SS$^+$	41.8	45.7	11.7
	CRO	**36.9**	**37.0**	0.4
3	h-BFOA	41.2	5324.0	11736.8
	SS$^+$	41.9	47.0	13.7
	CRO	**41.0**	**41.1**	0.4
4	h-BFOA	32.7	32.9	0.6
	SS$^+$	32.8	32.9	0.4
	CRO	**32.6**	**32.7**	0.1
5	h-BFOA	51.5	8605.2	17592.9
	SS$^+$	51.5	59.4	20.8
	CRO	**51.4**	**51.4**	0.3
6	h-BFOA	**44.0**	46.0	8.9
	SS$^+$	46.5	47.5	3.1
	CRO	**44.0**	**43.9**	0.2
7	h-BFOA	**56.1**	2451.8	7190.3
	SS$^+$	56.7	65.1	21.2
	CRO	**56.1**	**56.2**	0.5
8	h-BFOA	**44.3**	44.5	0.6
	SS$^+$	**44.3**	44.7	1.7
	CRO	44.4	**44.4**	0.3
9	h-BFOA	**53.0**	53.6	2.7
	SS$^+$	52.9	59.5	27.8
	CRO	53.1	**53.1**	0.3
10	h-BFOA	**46.7**	742.1	3734.6
	SS$^+$	49.9	52.3	5.3
	CRO	**46.7**	**46.7**	0.2
11	h-BFOA	57.7	3641.9	9149.9
	SS$^+$	**57.6**	60.6	7.7
	CRO	**57.6**	**57.7**	0.4
12	h-BFOA	**47.2**	47.3	0.5
	SS$^+$	46.9	**47.1**	0.7
	CRO	**47.2**	47.2	0.2
13	h-BFOA	**35.2**	7760.7	17503.4
	SS$^+$	35.6	37.1	7.3
	CRO	**35.2**	**35.2**	0.2
14	h-BFOA	**31.0**	**31.1**	0.5
	SS$^+$	33.4	34.3	4.6
	CRO	**31.0**	**31.1**	0.3
15	h-BFOA	**40.4**	44.4	23.1
	SS$^+$	40.7	67.2	117.9
	CRO	**40.4**	**40.5**	0.4
16	h-BFOA	**29.2**	29.7	1.8
	SS$^+$	**29.2**	**29.3**	0.4
	CRO	**29.2**	**29.3**	0.2

minutes than SS$^+$ and h-BFOA in 20 minutes (see the results in [9]), and achieving the lower standard deviation results.

Regarding the statistical analysis for median MSE values, the result of applying Friedman's test is $\chi^2_F = 6.125$, and the corresponding p-value is 0.047. Given that the p-value is lower than the considered level of significance ($\alpha = 0.05$), the test concludes that there are significant differences among the results. However, the results obtained for Bonferroni-Dunn's test and Holm's test (Table VI) reveal that only the differences between the control method (CRO) and SS$^+$ are significant. As to mean MSE values, the statistical significance is higher, with $\chi^2_F = 21.875$ and a p-value ($1.77 \cdot 10^{-05}$) under the given α. In this case, the statistical tests (Table VII) also confirms significant differences between CRO and both h-BFOA and SS$^+$.

Table VI. RANKING OBTAINED THROUGH FRIEDMAN'S TEST AND STATISTICAL P-VALUES WITH CRO AS CONTROL METHOD FOR HOLM'S TEST ACCORDING TO THE MEDIAN MSE VALUE.

Algorithm	Mean Rank	Bonferroni-Dunn ρ	Holm ρ
CRO	1.68	-	-
h-BFOA	1.81	1	1
SS$^+$	2.50	0,04	0,02

Table VII. RANKING OBTAINED THROUGH FRIEDMAN'S TEST AND STATISTICAL P-VALUES WITH CRO AS CONTROL METHOD FOR HOLM'S TEST ACCORDING TO THE MEAN MSE VALUE.

Algorithm	Mean Rank	Bonferroni-Dunn ρ	Holm ρ
CRO	1.06	-	-
SS$^+$	2.31	$< 10^{-4}$	$< 10^{-4}$
h-BFOA	2.63	$< 10^{-5}$	$< 10^{-6}$

V. CONCLUSIONS

In this work, we have faced the problem of intensity-based medical IR, a classic and complex problem with the aim of obtaining precise and robust solutions efficiently. Until now, the former goal had been accomplished, but in an inefficient way, as the considered MHs needed a long time to converge. In order to overcome this situation, we have considered the use of a novel evolutionary-type approach called CRO algorithm. This algorithm is based on the natural processes taking place within a coral reef, such as reef formation and expansion, coral reproduction and fight for a space in the colony. We adapted and applied CRO to the IR problem, following an *intensity-based* approach. Altogether, the intrinsic capabilities of CRO, such as the solid balance between diversity and specificity, combined with the particular augmentations of the intensity-based IR approach, i.e. multiresolution and restart mechanism, provide strong indications for a precise, robust and efficient optimization algorithm.

We followed the experimental guidelines of [9] to assess the performance of our proposal. We considered sixteen IR scenarios using pairs of four different brain MRI images, with different levels of noise and sclerosis lesions. These images were extracted from the BrainWeb database at McGill University [11]. In order to evaluate the performance of our proposal, we compared CRO with the best performing IR methods of our past study measuring median and mean values of the MSE similarity metric. CRO provided outstanding results considering both values of this metric in most of the scenarios, thus demonstrating the efficiency, robustness and suitability of the proposed algorithm when tackling complex optimization problems such as *intensity-based* medical IR.

As future work, we aim to extend our study and examine the behavior of CRO in other registration tasks, i.e., considering *feature-based* IR problems. In addition, it may be interesting to further improve the performance of CRO. One of the possible options is to modify its original design and consider a CRO variant with substrate layers. This variant integrates different exploration operators to establish a competitive co-evolution process.

ACKNOWLEDGMENTS

This work has been partially supported by the projects TIN2015-67661-P, including European Regional Development Funds (ERDF), from the Spanish Ministry of Economy and

TIN2014-54583-C2-2-R from the Spanish Ministerial Commission of Science and Technology (MICYT).

REFERENCES

[1] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computating*, vol. 21, pp. 977–1000, 2003.

[2] X. Y. Wang, S. Eberl, M. Fulham, S. Som, and D. D. Feng, "Data registration and fusion," in *Biomedical information technology*, D. D. Feng, Ed. Burlington, USA: Academic Press, 2008, pp. 187–210.

[3] A. Goshtasby, *2D and 3D Image Registration*. Wiley Interscience, 2005.

[4] P. J. Besl and N. D. McKay, "A method for registration of 3D shapes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 239–256, 1992.

[5] S. Klein, M. Staring, and J. P. W. Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines," *IEEE Trans. on Image Processing*, vol. 16, no. 12, pp. 2879–2890, 2007.

[6] S. Klein, J. Pluim, M. Staring, and M. Viergever, "Adaptive stochastic gradient descent optimisation for image registration," *International Journal of Computer Vision*, vol. 81, pp. 227–239, 2009.

[7] F. Glover and G. A. Kochenberger, Eds., *Handbook of Metaheuristics*. Boston, USA: Kluwer Academic Publishers, 2003.

[8] A. Valsecchi, S. Damas, J. Santamaría, and L. Marrakchi-Kacem, "Intensity-based image registration using scatter search," *Artificial Intelligence in Medicine*, vol. 60, no. 3, pp. 151–163, 2014.

[9] E. Bermejo, A. Valsecchi, S. Damas, and O. Cordón, "Bacterial foraging optimization for intensity-based medical image registration," in *2015 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2015, pp. 2436–2443.

[10] S. Salcedo-Sanz, J. Del Ser, I. Landa-Torres, S. Gil-López, and J. A. Portilla-Figueras, "The coral reefs optimization algorithm: A novel metaheuristic for efficiently solving optimization problems," *The Scientific World Journal*, vol. 2014, p. 15, 2014.

[11] D. L. Collins, A. P. Zijdenbos, V. Kollkian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. on Medical Imaging*, vol. 17, pp. 463–468, 1998.

[12] M. Svedlow, C. D. Mc-Gillem, and P. E. Anuta, "Experimental examination of similarity measures and preprocessing methods used for image registration," in *Symposium on Machine Processing of Remotely Sensed Data*, P. Swain, D. Morrison, and D. Parks, Eds., vol. 4(A), Indiana, USA, 1976, pp. 9–17.

[13] M. A. Audette, F. P. Ferrie, and T. M. Peters, "An algorithmic overview of surface registration techniques for medical imaging," *Medical Image Analysis*, vol. 4, no. 3, pp. 201–217, 2000.

[14] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Trans. on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.

[15] F. Maes, D. Vandermeulen, and P. Suetens, "Comparative evaluation of multiresolution optimization strategies for image registration by maximization of mutual information," *Medical Image Analysis*, vol. 3, no. 4, pp. 373–386, 1999.

[16] S. Klein, M. Staring, and J. P. W. Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines," *IEEE Trans. on Image Processing*, vol. 16, no. 12, pp. 2879–2890, 2007.

[17] J. M. Rouet, J. J. Jacq, and C. Roux, "Genetic algorithms for a robust 3-D MR-CT registration," *IEEE Trans. on Information Technology in Biomedicine*, vol. 4, no. 2, pp. 126–136, 2000.

[18] S. M. Yamany, M. N. Ahmed, and A. A. Farag, "A new genetic-based technique for matching 3D curves and surfaces," *Pattern Recognition*, vol. 32, pp. 1817–1820, 1999.

[19] P. Chalermwat, T. El-Ghazawi, and J. LeMoigne, "2-phase GA-based image registration on parallel clusters," *Future Generation Computer Systems*, vol. 17, pp. 467–476, 2001.

[20] S. Zambanini, R. Sablatnig, H. Maier, and G. d. Langs, "Automatic image-based assessment of lesion development during hemangioma follow-up examinations," *Artificial Intelligence in Medicine*, vol. 50, no. 2, pp. 83–94, 2010.

[21] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International Journal of Computer Vision*, vol. 13, no. 2, pp. 119–152, 1994.

[22] E. Bermejo, O. Cordón, S. Damas, and J. Santamaría, "A comparative study on the application of advanced bacterial foraging models to image registration," *Information Sciences*, vol. 295, pp. 160–181, 2015.

[23] K. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," *Control Systems, IEEE*, vol. 22, no. 3, pp. 52–67, Jun 2002.

[24] H. G. Beyer and K. Deb, "On self-adaptive features in real-parameter evolutionary algorithms," *IEEE Trans. on Evolutionary Computation*, vol. 5, no. 3, pp. 250–270, June 2001.

[25] F. Glover, "Heuristic for integer programming using surrogate constraints," *Decision Sciences*, vol. 8, pp. 156–166, 1977.

[26] F. Glover, M. Laguna, and R. Martí, "Scatter search," in *Advances in Evolutionary Computation: Theory and Applications*, A. Ghosh and S. Tsutsui, Eds. New York: Springer-Verlag, 2003, pp. 519–537.

[27] M. Lozano, F. Herrera, N. Krasnogor, and D. Molina, "Real-coded memetic algorithms with crossover hill-climbing," *Evolutionary Computation*, vol. 12, no. 3, pp. 273–302, 2004.

[28] S. Salcedo-Sanz, "A review on the coral reefs optimization algorithm: new development lines and current applications," *Progress in Artificial Intelligence*, vol. in press, pp. 1–15, 2017.

[29] E. Bermejo, O. Cordón, S. Damas, and J. Santamaría, "Quality time-of-flight range imaging for feature-based registration using bacterial foraging," *Applied Soft Computing*, vol. 13, no. 6, pp. 3178–3189, 2013.

[30] P. Rogelj and S. Kovacic, "Validation of a Non-Rigid Registration Algorithm for Multimodal Data," in *SPIE in Medical Imaging*, M. Sonka and J. M. Fitzpatrick, Eds., San Diego, USA, 2002, pp. 299–307.

[31] O. Monga, S. Benayoun, and O. Faugeras, "From partial derivatives of 3D density images to ridges lines," in *Computer Vision and Pattern Recognition*. Champaign, Illinois, USA: IEEE, 1992, pp. 354–389.

[32] A. Valsecchi, S. Damas, J. Santamaría, and L. Marrakchi-Kacem, "Genetic algorithms for voxel-based medical image registration," in *Computational Intelligence in Medical Imaging (CIMI), 2013 IEEE Fourth International Workshop on*, April 2013, pp. 22–29.

[33] M. Friedman, "A comparison of alternative tests of significance for the problem of m rankings," *The Annals of Mathematical Statistics*, vol. 11, no. 1, pp. 86–92, 1940.

[34] O. J. Dunn, "Multiple Comparisons Among Means," *Journal of the American Statistical Association*, vol. 56, no. 293, pp. 52–64, mar 1961.

[35] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1979.