Threshold selection, mitosis and dual mutation in cooperative co-evolution: application to medical 3d tomography

Franck Vidal, Evelyne Lutton, Jean Louchet, Jean-Marie Rocchisani

Threshold selection, mitosis and dual mutation in cooperative co-evolution: application to medical 3D tomography

Franck P. Vidal[1], Evelyne Lutton[2], Jean Louchet[3], and Jean-Marie Rocchisani[4]

[1] Department of Radiation Oncology, University of California, San Diego, CA.
franck.p.vidal@gmail.com
[2] INRIA - Saclay-Île-de-France, AVIZ team, Orsay, France.
Evelyne.Lutton@inria.fr
[3] Artenia, Châtillon, France.
Jean.Louchet@gmail.com
[4] Paris XIII University, UFR SMBH & Avicenne hospital, Bobigny, France.
jean-marie.rocchisani@avc.aphp.fr

Abstract. We present and analyse the behaviour of specialised operators designed for cooperative coevolution strategy in the framework of 3D tomographic PET reconstruction. The basis is a simple cooperative co-evolution scheme (the "fly algorithm"), which embeds the searched solution in the whole population, letting each individual be only a part of the solution. An individual, or fly, is a 3D point that emits positrons. Using a cooperative co-evolution scheme to optimize the position of positrons, the population of flies evolves so that the data estimated from flies matches measured data. The final population approximates the radioactivity concentration. In this paper, three operators are proposed, threshold selection, mitosis and dual mutation, and their impact on the algorithm efficiency is experimentally analysed on a controlled test-case. Their extension to other cooperative co-evolution schemes is discussed.

1 Introduction

Evolutionary algorithms have been proven efficient to solve the inverse problem of 3D data reconstruction in tomography [1], and particularly of positron emission tomography (PET) reconstruction in nuclear medicine [2–4].

In PET, a positron emitter is used as radionuclide for labelling. Positrons generally lead to an annihilation reaction, that emits two photons of 511 keV in opposite directions. This radiation is detected in coincidence, i.e. using the difference in arrival times of the detected photons of each pair, and considering that each annihilation produces two photons emitted in exactly opposite directions. The line joining the detectors that have been activated for a given pair of photons is called "line of response" (LOR). An overview of reconstruction methods in nuclear medicine can be found in [5].

In previous work, we showed that a cooperative coevolution strategy (or Parisian evolution) called "flies algorithm" [6] could be used in Single-Photon

Emission Computed Tomography (SPECT) reconstruction [1], and also in PET reconstruction in 2D-mode [2, 3], and in Fully-3D-mode [4]. The marginal fitness was used to propose new operators, i) the threshold selection, and ii) the mitosis, but no performance analysis of these operators has been performed so far.

This paper addresses this deficiency and it analyses the impact of each operator on the performance of a PET reconstruction algorithm on a controlled test-case. A new operator (namely the dual mutation) and a pre-initialisation of the flies' position using back-projection are also described and analysed. Standard PET reconstruction algorithms are reviewed in Section 2. It is followed by an overview of the fly algorithm for PET reconstruction. The three operators that control a varying population size scheme are presented in Section 4, as well as an alternate initialisation process. Experimental setup and analysis are given in Section 5, before presenting some conclusions and future work in Section 6.

2 Standard PET reconstruction algorithms

Tomography reconstruction algorithms can be divided into two main categories.

On the one hand, there are analytical methods. These are based on a continuous modelling and the reconstruction consists in the inversion of measurement equations, such as the well known Filtered Back-Projection (FBP). This method is now rarely used due to strong artefacts in the reconstructed data (see Fig. 5) and also because the correction of imaging physics effects need to be undertaken before the reconstruction, leading to a systematic positive bias in the reconstructed volume.

On the other hand, there are iterative methods. This class of methods can be split into two kinds. Algebraic methods are used in X-ray Computed Tomography (CT); statistical methods are used in nuclear medicine for both SPECT and PET [7]. They take into account noise, and the correction of imaging physics can be applied during the reconstruction in the iterative steps. Iterative methods are relatively easy to model. In practice, the volume is usually discretised into voxels. Each voxel intensity is treated as an unknown. A system of linear equations is defined according to the imaging geometry and physics: $\mathbf{p} = \mathbf{R}\,\mathbf{f}$, with $\mathbf{f}$ the volume to recover, $\mathbf{p}$ the measured data, $\mathbf{R}$ the system model. Imaging physics, such as non-uniform attenuation, scatter, etc. can be modelled in $\mathbf{R}$, whereas they are difficult to handle in an analytic algorithm. The system of equations is finally solved using the iterative algorithm.

There are different ways to implement these iterative methods. The main differences are about the computation of the projections, the physics corrections (scattering, random, attenuation, etc.) are applied, and how the error corrections are applied in the estimated projections.

The Maximum Likelihood - Expectation Maximisation (ML-EM) (or 'EM") is a common algorithm in SPECT and PET. It assumes Poisson noise is present in the projection data. ML-EM does not produce artefacts seen in FBP reconstructions, and it has a better signal-to-noise ratio in region of low concentration. However, the algorithm converges slowly.

The Ordered Subset - Expectation Maximization (OS-EM) has been proposed to speed-up convergence of the EM algorithm. Its principle is to reduce the number of projections used at each iteration of the EM algorithm. Projections are grouped in K sub-groups. The projections of a sub-group are uniformly distributed around the volume to reconstruct.

3 PET reconstruction using the fly algorithm

The fly algorithm for tomography reconstruction follows the iterative paradigm. The steps of the iterative method can be described as follows:

1. Each individual, or fly, corresponds to a 3D point. Initially, the flies' position is randomly generated in the volume within the scanner. The population of flies corresponds to the tracer density in the patient.
2. To produce estimated projection data, each fly mimics a radioactive emitter, i.e. a stochastic simulation of annihilation events is performed. For each annihilation event, a photon is emitted in a random direction. A second photon is then emitted in the opposite direction. If both photons are detected by the scanner, the corresponding LOR is recorded. The scanner properties (e.g. detector blocks and crystals positions) are modelled, and each fly is producing an adjustable number of annihilation events.
3. The optimisation is performed using genetic operations. The fitness function used during the selection operation takes into account the comparison between the estimated projections and the measured projections.
4. Using genetic operations to optimise the position of radioactive emitters, the population of flies evolves so that the population total pattern matches measured data.
5. Instead of a "generational" evolutionary strategy, in which at each loop every individual (fly) will be eliminated and replaced with a new fly, we chose a "steady state" evolutionary strategy.

Note that in classical evolutionary approaches, each individual in the population is a potential solution; in the Fly approach, a subset of the evolving population itself is the representation of the solution. After convergence, the "good" flies (see Section 4.1) are then extracted to form the reconstructed volume.

4 Varying population size scheme in a cooperative co-evolution algorithm

Cooperative co-evolution strategies rely on a "social" formulation of the optimisation problem, where individuals collaborate or compete in order to collectively build a solution. The fly algorithm is a mono-population strategy (Parisian approach): all flies contribute independently and collectively to build the solution. In [8] a variable sized population Parisian GP strategy has been successfully used on a cooperative co-evolution, based on adaptive population deflating and inflating schemes. We test in this paper an "inflating-only" strategy, the mitosis, described below, to gradually increase the precision of the reconstructed data.

4.1 Marginal fitness

In this application, the similarities or discrepancies between the estimated projection data and the measured projection data provided by the imaging system have to be assessed. We chose City Block distance (also known as Manhattan distance) as the fitness metrics to measure the distance between two LOR sets. It provides a good compromise between speed and accuracy. Eq. 1 provides the global fitness, i.e. the population's cost:

$$\mathrm{dist}(LOR_m, LOR_e) = \sum_{i}^{M} \sum_{j}^{M} |LOR_m(i,j) - LOR_e(i,j)| \tag{1}$$

with $\mathrm{dist}(LOR_m, LOR_e)$ the City Block distance between LOR_m and LOR_e, the set of LORs for the measured data and the estimated data respectively, $LOR(i,j)$ is the number of counts of a LOR between the photon detectors i and j, M is the total number of photon detectors within the imaging system. LOR sets are efficiently implemented using triangular sparse matrices to reduce the amount of memory needed to store the data. The smaller global fitness is, the closer the simulated data will be to the measured data.

In [1], we showed that, when we were addressing the SPECT problem, if we defined the fitness of a fly as the consistency of the image pattern it generates, with the actual images, it gave an important bias to the algorithm with a tendency of the smaller objects to disappear. To address this, we introduced marginal evaluation to assess a given fly. We use a similar approach in PET:

$$\mathrm{F_m}(x) = \mathrm{dist}\left(LOR_e - \{LOR_x\}, LOR_m\right) - \mathrm{dist}\left(LOR_e, LOR_m\right) \tag{2}$$

with $\mathrm{F_m}(x)$ the marginal fitness of Fly x, and $LOR_e - \{LOR_x\}$ is the set of LORs simulated by the whole population without Fly x. In practice, each fly needs to keep a record of its simulated LORs.

The fitness of a given fly will only be positive when the global cost is lower (better) in presence rather than in the absence of this fly.

4.2 Threshold selection

The fly to be killed is randomly chosen by the "selection" operator, with a bias towards killing "bad" individuals. On the other hand, if the new fly is to be created by mutation of another fly, this fly is randomly chosen by the "selection" operator, with a bias towards reproducing "good" individuals. Classical selection operators are ranking, roulette wheel and tournament [9]. In our algorithm, as each fly's fitness is the value of its (negative or positive) contribution to the quality of the whole population, we managed to simplify and speed up the selection process by using a fixed fitness threshold. Any "bad" fly (its fitness is negative) is a candidate for death, and any "good" fly (its fitness is positive) is a candidate for mutation.

4.3 Mitosis

When the number of flies with a negative fitness decreases, the threshold selection fails to provide flies to be killed in an acceptable time. It also means that the reconstruction is optimum at the current resolution. If the resolution is acceptable, i.e. there are enough flies to approximate the radio-tracer concentration, then the algorithm can stop and the reconstructed volume is extracted using flies with a positive fitness. If not, a mitosis operator is triggered to gradually increase the population size. Each fly is split into two new flies to double the population size. One of the two flies is then mutated.

4.4 Dual mutation

To optimise the flies' position, our algorithm takes advantage of a mutation operator. When a new fly (b) is created by mutation of an old "good" fly (a), the position of Fly b is first initialised to the same position as Fly a. The new fly is then stochastically translated in any direction, and LORs are randomly generated from that fly. The length of the translation vector is a random variable that follows a Gaussian law whose mutation variance is σ^2. It needs first to be set to a large value to better explore the search space. However, a constant large mutation variance will lead to blurred reconstructed volumes. σ has therefore to be gradually reduced.

The use of adaptive mutations in evolutionary algorithms is an ancient idea, directly inspired by natural adaptive phenomena, e.g. mutations simulated by stress [10]. In artificial evolution, various adaptive schemes have been considered for mutation [11], depending of the parameter to be adapted (standard deviation, σ [12], or mutation law [13] for continuous mutation, mutation probability for discrete mutations [14]). Regarding the adaptation of σ, one can distinguish several strategies :

- σ is directly adapted to local measurements, like fitness [15] or local regularity [16],
- σ is tuned depending on some success measurement: in this category fall the famous $1/5^{th}$ rule proposed by Schewefel [17, 18],
- σ is subject to an adaptive pressure itself, it is self-adapted [19]: σ is considered as an additional parameter in the genome, and a log-normal Gaussian law is used to control the "mutation over the mutation".

These techniques have been proven efficient in various cases, depending on the fitness function and the genetic engine. It has however to be noticed that the sophistication of a mutation operator has a computational cost, and that some very rough schemes may perform better due to their capability to rapidly test numerous sample points [20].

Concurrent testing with various subpopulation has been also considered for mutation law adaptation [13].

Here we propose an adaptive mutation scheme based on the concurrent testing of two alternative σ values (σ_{low} and σ_{high}, with $k\ \sigma_{low} = \sigma_{high}$). The update

rule is multiplicative as for the $1/5^{th}$ rule. If σ_{high} gives the best results during the previous period, then both mutation variances are increased by a predefined factor (pf, with $pf > 1$). If σ_{low} gives better results, then the variances are multiplied by $\frac{1}{pf}$. The major advantage of this dual mutation scheme is to provide a fully automatic method to adapt the mutation variance, whilst keeping the administration cost of the algorithm relatively light. Additionnally, this scheme does not need to make any assumption on the ideal success rate of the mutation as in the $1/5^{th}$ rule. In practice, the global fitness is recorded after each mutation. The cumulative difference of the global fitness ($\Delta(\sigma)$) before and after mutations is computed to determine which σ value provides the best performance over a given period of time. To prevent oscillation of σ values, a criteria can be added to avoid changes when both σ_{high} and σ_{low} provide relatively similar results, e.g. when the absolute difference between $\Delta(\sigma_{low})$ and $\Delta(\sigma_{high})$, relative to the current global fitness, is below a given threshold (t_{mut}).

4.5 Initialisation of flies on LORs

Iterative reconstruction methods generally make use of a constant volume as an initial estimate of the volume (see Fig. 5(a)).

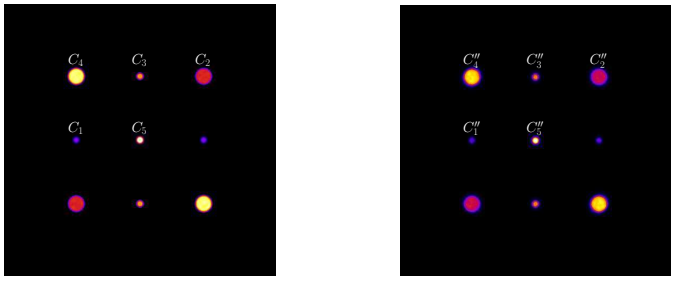
However, to speed-up the reconstruction process, a volume is first reconstructed using a fast analytical algorithm, the simple back-projection, that we implemented on the graphics card using OpenGL. The algorithm consists in back-projecting each LOR into the volume space. Pixels along the path of a LOR are updated uniformly. This operation is fast and provides the evolutionary algorithm with an initial guess of the volume (see Fig. 5(c)). For each voxel of the initial estimate, a given number of flies is assigned depending on the voxel intensity (see Fig. 5(b)).

5 Results

The validity of the reconstruction method has been addressed in [4]. In this paper, we focus on the evaluation of the performance of the new genetic operators. For each test case, 750000 new individuals have been created. For each tested configuration, the reconstruction has been repeated 20 times, and the final global fitness was recorded. For every test, unless specified, the dual variance and the threshold selection operators have been enabled. Results are presented using box plots (also called box-and-whisker diagrams).

5.1 Experimental setup

Here, a single ring PET system is considered. Its radius is about 430mm. The ring is made of 72 linear blocks that include 8 crystals each. The width of a crystal is about 4.5mm. Fig. 1(a) shows the reference image. It includes nine cylinders having two different radii (1 cm and 2.5 cm) and five different radioactivity concentrations ($C_1 = 114,590$ count/ml, $C_2 = 2C_1$, $C_3 = 3C_1$, etc.)



(a) Reference image. (b) Reconstructed image.

Fig. 1. Slices (512 × 512 pixels) through the cylinders.

5.2 Threshold selection

The size of the population is fixed (160000 flies), i.e. no mitosis has been triggered. The performance of the threshold selection and the tournament selection are presented in Fig. 2. Both operators provide similar performance. The

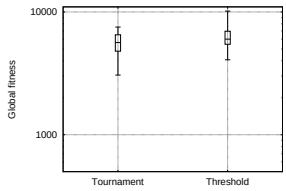


Fig. 2. Performance of the threshold selection and of the tournament selection.

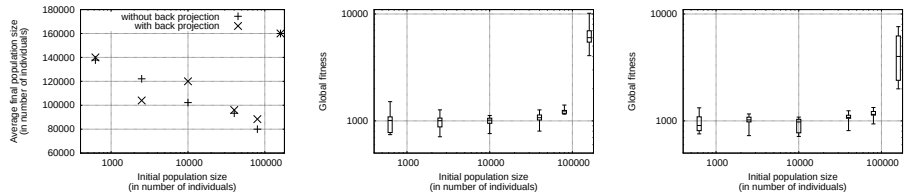
threshold selection is then preferred because of the additional information that it brings: enable mitosis, and provide a convergence criteria at a given resolution (i.e. for a given size of population).

5.3 Mitosis

Two variables have to be assessed at the end of the reconstruction: the current size of the population, and the global fitness. The larger the final population, the better the image resolution can be obtained. The smaller the global fitness, the closer the estimated data to the measured data is.

Fig. 3(a) shows the average number of flies in the final population depending on the initial size of the population (625, 2500, 10000, 40000, 80000, and 160000 flies). When the size of the population is 160000 flies, no mitosis has been triggered. Fig. 3(b) shows the corresponding global fitness.

Similar performance in term of global fitness is obtained when the initial population size is below 10000 flies. The highest final population size is obtained



(a) Average final population size.

(b) Mitosis without initialisation of the flies' position.

(c) Mitosis with initialisation of the flies' position using Fig. 5(c)

Fig. 3. Performance of the mitosis operator using variable initial population sizes.

with the smallest initial population size. Then, the reconstruction converges much faster when the initial population is small (smaller global fitness and bigger final population size). These results validate the efficiency of the mitosis operator.

5.4 Dual mutation

The initial σ_{low} value in this test is 35mm, pf is equal to $\sqrt[3]{2}$, and σ_{high} is equal to $\sqrt[3]{2}\sigma_{low}$. Different threshold values (t_{mut}) have been tested to limit oscillations of σ values (Fig. 4(a)). Larger values not only prevent oscillations, they also

(a) When the initial population size is 625 flies, and depending on t_{mut} value.

(b) When the initial population size is 160000 flies with: (1) constant σ = 0.1mm, (2) constant σ = 0.01mm, (3) t_{mut} = 0.0%, and (4) t_{mut} = 0.05%.

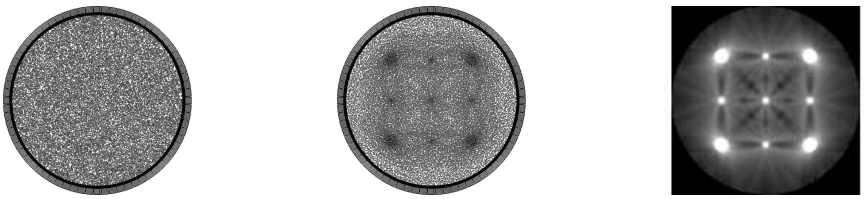
Fig. 4. Performance of the dual variance operator.

prevent any change of σ values, leading to unsatisfactory results.

Fig. 4(b) shows the global fitness obtained i) using a constant variance (see (1) and (2)), or ii) enabling dual mutation operator (see (3) and (4)). The best results are observed using the dual mutation operator with a very low t_{mut} value.

5.5 Initialisation of flies on LORs

Fig. 5(a) and Fig. 5(b) show two possible initial estimates of the radio-active concentration. In the first case, flies are uniformly located within the space in



(a) Uniform distribution of the flies.

(b) Initialisation of the flies' position using Fig. 5(c)

(c) Image reconstructed using the simple back projection algorithm.

Fig. 5. Initial estimates of the reconstructed image.

the imaging system. In the latter case, the position of flies is initialised using the simple back projection. Fig. 3 shows the performance of both strategies when the mitosis operator is enabled.

When the initial size of the population is relatively large, the algorithm converges much faster using this initialisation step. This is not the case when the initial size of the population is relatively small. It may be due to the fact that the algorithm converges fast enough when only a few flies are used. When the initial number of flies is slightly higher, the reconstruction converges faster when the position of flies are initialised using the back projection.

6 Conclusion and futher works

We have presented new operators in cooperative co-evolution and validated their efficiency using a controlled test-case in PET reconstruction. Both the threshold selection, mitosis and dual mutation operators have shown their usefulness and ability. Experimental statistics show that the threshold selection perform as well as the tournament selection, but it has the great advantage of bringing a convergence criterion related to the current resolution. Additionally, it allows to trigger an automatic mitosis, i.e. doubling the population size, to improve the resolution. Best performance, both in term of final resolution and convergence, are obtained using small initial population size. The dual mutation operator provides an adaptive mutation variance that has proven to be better than using fixed mutation variances.

Such operators can be used in any other cooperative co-evolution schemes as soon as a marginal fitness can be considered as beneficial, that obviously depends on the computation cost of the marginal fitness. For instance, threshold selection, mitosis and dual selection will be considered as further work for the original fly algorithm on a stereo-vision application ([6]). The marginal fitness will also be considered for developing a "deflating operator". This additional mechanism for controlling the population size may be interesting in the case of applications whose resolution does not depend on the size of the final population.

Further work will also include the correction of photon attenuation and Compton scattering, and a concurrent study with the OS-EM algorithm.

References

1. Bousquet, A., Louchet, J., Rocchisani, J.M.: Fully three-dimensional tomographic evolutionary reconstruction in nuclear medicine. In: Proc. EA'07. Volume 4926 of LNCS. (2007) 231–242
2. Vidal, F.P., Lazaro-Ponthus, D., Legoupil, S., Louchet, J., Lutton, E., Rocchisani, J.: Artificial evolution for 3D PET reconstruction. In: Proc. EA'09. Volume 5975 of LNCS., Springer (2009) 37–48
3. Vidal, F.P., Louchet, J., Lutton, E., Rocchisani, J.M.: PET reconstruction using a cooperative coevolution strategy in LOR space. In: IEEE Nuclear Science Symposium Conference Record, Orlando, Florida, IEEE (October 2009) 3363–3366
4. Vidal, F.P., Louchet, J., Rocchisani, J.M., Lutton, E.: New genetic operators in the fly algorithm: application to medical PET image reconstruction. In: EvoApplications 2010. Volume 6024, Part I of LNCS., Springer (2010) 292–301
5. Zeng, G.L.: Image reconstruction – a tutorial. Comput. Med. Imaging Graph. **25**(2) (2001) 97–103
6. Louchet, J.: Stereo analysis using individual evolution strategy. In: Proc. ICPR'00. (2000) 1908
7. Vandenberghe, S., D'Asseler, Y., Van de Walle, R., Kauppinen, T., Koole, M., Bouwens, L., Van Laere, K., Lemahieu, I., Dierckx, R.A.: Iterative reconstruction algorithms in nuclear medicine. Comput. Med. Imaging Graph. **25** (2001) 105–111
8. Barriere, O., Lutton, E.: Experimental analysis of a variable size mono-population cooperative-coevolution strategy. In: Proc. NICSO'08. (2008)
9. Baeck, T., Fogel, D.B., Michalewicz, Z., eds.: Evolutionary Computation 1: Basic Algorithms and Operators. Taylor & Francis (2000)
10. Rosenberg, S.M.: Evolving responsively: adaptive mutation. Nat. Rev. Genet. **2** (2001) 504–515
11. Eiben, A.E., Hinterding, R., Michalewicz, Z.: Parameter control in evolutionary algorithms. IEEE T. Evolut. Comput. **3**(2) (1999) 124–141
12. Bäck, T., Schwefel, H.P.: An overview of evolutionary algorithms for parameter optimization. Evol. Comput. **1**(1) (1993) 1–23
13. Chellapilla, K.: Combining mutation operators in evolutionary programming. IEEE T. Evolut. Comput. **2**(3) (1998) 91–96
14. Ochoa, G.: Setting the mutation rate: Scope and limitations of the 1/L heuristic. In: Proc. GECCO'02, Morgan Kaufmann Publishers Inc. (2002) 495–502
15. Lis, J.M., Orlowska-Kowalska, T.: Application of evolutionary algorithms with adaptive mutation to the identification of induction motor parameters at standstill. In: Proc. EUROCON'07. (2007) 1786–1791
16. Lutton, E., Lévy Véhel, J.: Pointwise regularity of fitness landscapes and the performance of a simple ES. In: Proc. CEC'06. (2006) 16–21
17. Schwefel, H.P.: Numerical optimization of computer models. Wiley (1981)
18. Beyer, H.G., Schwefel, H.P.: Evolution strategies - a comprehensive introduction. Nat. Comput. **1**(1) (2002) 3–52
19. Bäck, T.: Self-adaptation in genetic algorithms. In: Proceedings of the First European Conference on Artificial Life, MIT Press (1992) 263–271
20. Collet, P., Lutton, E., Louchet, J.: Issues on the optimisation of evolutionary algorithm code. In: Proc. CEC'02. (2002)