Tuning Genetic Algorithms for Underwater Acoustics Using *a priori* Statistical Information

M. João Rendas

I3S, CNRS URA 1376
Univ. Nice Sophia Antipolis, France.
rendas@mimosa.unice.fr

Georges Bienvenu

THOMSON Marconi Sonar
BP. 157, 06903 Sophia Antipolis Cedex
Sophia Antipolis, France.

ABSTRACT

In this paper we present a new technique for the evaluation/selection procedures of genetic algorithms, to be used in the context of parameter estimation problems. The proposed algorithm uses *a priori* information about the structure of the surface of which an extremum is being searched. For parameter estimation problems, the availability, at each iteration of a genetic algorithm, of a collection of samples of the ambiguity surface of the problem, enables the determination of the correlation between the *observed* ambiguity surface (at the sampled points) and the *predicted* ambiguity surface. The consideration of this information allows early detection of secondary extrema (which yield an ambiguity surface which does not correlate well with the observed one) and thus contributes to speed the convergence of the algorithm to the global optimal values. The paper applies the proposed technique to a source localization problem.

1. INTRODUCTION

Genetic algorithms, originally proposed in [5] are a powerful optimization tool for underwater acoutics applications, which have shown their interest in connection with matched field techniques [4, 8, 9]. Genetic Algorithms (GA's) are generic stochastic search techniques which, under mild conditions on the reproduction operators and on the objective function, guarantee an asymptotic convergence to the global optimum even under the existence of secondary extrema or discontinuities [10]. Requiring *no a priori information* on the shape and characteristics of the problem being solved, they are a general-purpose technique, unsensitive to the peculiarities of the application being considered.

The price paid for the robustness of GA's is the slow convergence towards the optimal solution being sought. It has been observed, [1], that other variants of Evolutionary Algorithms, such as Evolutionary Programming (EP) and, more specially, Evolutionary Strategy (ES), yield much faster convergence rates. This improvement is obtained by using more sophisticated genetic operators, which give ES's and EP's the ability to self-adapt to the characteristics of the objective function.

In this paper, we show how *a priori* information about the ambiguity structure [3, 6, 7] of the problem being solved can be used to speed-up the convergence of GA's. The key observation is that, for inverse problems in non-homogeneous mediums such as the ocean, the shape of the function being optimized depends on the true value of the unknown parameters being sought. The global shape of the ambiguity surface provides, in this manner, information about the true parameter values, and can thus be used to tailor the genetic operators and fitness criteria. Two major modifications are proposed:

- *in the evaluation/selection operator:* while canonical genetic algorithms for parameter optimization individually evaluate each element of the population, we present a novel evaluation procedure, which uses the values of the objective function on the whole population for evaluating each individual. Our method provides better rejection of secondary maxima, leading in this way to an increased efficiency of GA's;
- *in the reproduction operators:* while these are homogeneous in canonical GA's, i.e., independent of the sampled values of the objective function, we propose to adjust them using the information provided by the sampled points and the *a priori* model. In this way, the generation of new individuals is directed to regions that have higher probability of corresponding to the global optimum.

In this paper, only the implementation of the first modification is discussed, being presented in the context of coherent matched field passive source localization problems. We give in the next section a brief formal description of our approach.

2. COLLECTIVE EVALUATION

Let θ denote the complete set of unknown parameters in the source localization problem (which may include, for instance, parameters describing the radiated signal spectrum or environmental parameters for source focalization [2]). Let $\mathcal{A}(\theta : \theta_0)$ denote the problem ambiguity function, describing the resemblance of the probabilistic models corresponding to different parameter values.

Assume that at iteration k the population of the GA is the set of N_k points

$$P_k = \{\theta_1^k, \ldots, \theta_N^k\},$$

and let the function being optimized be $\mathcal{F}(\theta)$. Generation of the next population is based on the evaluation of $\mathcal{F}$ at all individuals of P_k:

$$\mathcal{F}^k = \{\mathcal{F}(\theta_1^k), \ldots, \mathcal{F}(\theta_N^k)\}.$$

Present GA algorithms determine the fitness of each individual θ_i^k, Φ_i^k, as a scaled positive version of $\mathcal{F}(\theta_i^k)$, $\Phi_i^k = \Phi(\mathcal{F}(\theta_i^k))$, and probabilistically select them for "reproducing" in the next generation, using their relative fitness:

$$p_s(\theta_i^k) = \frac{\Phi_i^k}{\sum_{j=1}^N \Phi_j^k}.$$

Assume now that the objective function is multi-modal, with important secondary lobes. The individuals falling in the secondary lobes of the objective function will be, with high probability (proportional to the importance of the secondary lobes), selected along with those falling in the main lobe of the ambiguity function. If a significant percentage of the population happens to fall in secondary lobes, or if their fitness is higher than those of the individuals inside the main lobe, selection of the next generation on the basis only of $\mathcal{F}^{\|}$ can direct the search towards the wrong regions in parameter space.

We propose a new definition of *collective fitness*, which makes the fitness of each individual θ_i^k dependent also on the value of the objective function for the rest of the population, i.e., on the entire vector $\mathcal{F}^k$ and not just on the value of $\mathcal{F}(\theta_i^k)$.

The new definition of the evaluation function for each individual θ_n^k is a measure of the *matching* between the sampled points and the predicted ambiguity surface. It can, for instance, be based on the likelihood function $L(\mathbf{r}|\theta_n^k)$ for the following estimation problem:

Given observations

$$\mathbf{r} = \{(\theta_1, \mathcal{F}_1), \ldots, (\theta_N, \mathcal{F}_N))\},$$

drawn from the known parametric family of probability distributions

$$p(\mathbf{r}|\theta^o) \sim g_{\theta_n^k}\left(\mathcal{F}_1 - \mathcal{A}(\theta_1 : \theta_n^k), \ldots, \mathcal{F}_N - \mathcal{A}(\theta_N : \theta_n^k)\right),$$

find an estimate of θ^o.

In the above problem, θ_n^k plays the role of the true value of the parameter being estimated, g_{θ^o} is a known distribution, parametrized by θ_n^k which describes the statistical deviation of the observed ambiguity from the predicted ambiguity, $\mathcal{A}(\theta : \theta_n^k)$, assuming that θ_n^k is the true value of the parameter. The above formulation assumes that the observed values of the objective function are distributed around those predicted by the ambiguity surface, with dispersion that may depend on the actual parameter value.

Using the overall shape of the observed ambiguity surface for selecting the next generation of GA's may effectively eliminate secondary extrema of the objective function. The following discussion explains the rational behind our approach. If the location of the secondary lobes predicted by statistical analysis of the problem does not match the observed local extrema of $\mathcal{F}$, then they must correspond to spurious extrema. For concreteness, assume that θ_{i_1} is the individual of the population to which it corresponds the largest value of the objective function:

$$\mathcal{F}(\theta_{i_1}) > \mathcal{F}(\theta_j), \forall j \neq i_1.$$

and let θ_{i_2} be the second best individual. Assume also that the ambiguity function for the problem is not symmetrical. Let $\mathcal{A}(\theta : \theta_{i_1})$ be the predicted ambiguity surface when the true parameter value is equal to θ_{i_1}, and $\mathcal{A}(\theta : \theta_{i_2})$ be the corresponding surface when $\theta^o = \theta_{i_2}$. If

$$\mathcal{A}(\theta_{i_2} : \theta_{i_1}) \simeq 0$$

but

$$\mathcal{A}(\theta_{i_1} : \theta_{i_2}) \simeq 1$$

one can conclude that θ_{i_1} is a secondary minimum, since it yields an ambiguity function that does not match the observed one. Implementation of this general procedure in the context of matched field source localization is described in section 4 of the paper.

3. TARGETED REPRODUCTION/MUTATION

The second way in which we propose to use *a priori* statistical information about the ambiguity structure of the problem concerns the definition of the stochastic operators that map one population into the next one:

$$R: \begin{array}{ccc} \Theta^N & \rightarrow & \Theta^N \\ P_k & \rightsquigarrow & P_{k+1} \end{array} .$$

R is traditionally a pre-defined isotropic probabilistic operator. In standard GA's, R does not depend neither on the iteration index k, nor on the characteristics of the population P_k to which they are applied.

We propose the use of generating operators that are built using the position of the extrema of the *collective fitness* defined above. More precisely, we claim that the new population should be obtained by sampling the following mixture distribution:

$$\begin{aligned} p(\theta^{k+1}|\mathbf{r}^k) &= \sum_{i=1}^{N} p(\mathbf{r}^k|\theta_i^k)\mathcal{A}'(\theta : \theta_i^k) \\ \mathbf{r}^k &= \{\mathbf{r}_i^k\}_{i=1}^N, \qquad \mathbf{r}_i^k = (\theta_i^k, \mathcal{F}(\theta_k^i)), \end{aligned}$$

where $\mathcal{A}'(\theta : \theta_i^k)$ is an unit-area version of the ambiguity surface, assuming that θ_i^k is the true value of the parameter.

4. APPLLICATION TO SOURCE LOCALIZATION

We apply in this section the proposed technique to a source localization problem. In the example presented, we considered the localization of a source in a channel with a bilinear velocity profile, exhibiting multipath. The axis of the duct occurs at a depth of 914 m, for a sound velocity of $v_{min} = 1480$ m/s, and the gradient of sound speed is $g_0 = -0335$ and $g_1 = -.013$ s^{-1} in the upper an lower layers, respectively. The receiving antenna is a vertical uniform array with 6 sensors and inter-element spacing of 15 m. The receiver searches for the maximum of the Bartlett spectrum,

$$\mathcal{F}(\theta) = \frac{1}{n_{freq}} \sum_{f=f_{min}}^{f_{max}} h_\theta^H C_f h_\theta ,$$

where C_f is the cross-spectral density matrix at frequency f, and h_θ is a normalized vector describing the nominal acoustic field received for parameters in vector θ. In our simulations, two frequencies $f = 50$ Hz and $f = 60$ Hz were used. We considered the use of focalization, by estimating the minimum sound velocity v_{min} along with the source range and depth, (R, D).

The ambiguity structure of the function being optimized can be computed by noting that the cross-spectral matrices C_f have the following structure:

$$C = S\, h_{\theta_0} h_{\theta_0}^H + \Sigma,$$

were θ_0 denotes the true values of the parameters being estimated, S is the signal power and Σ is the noise component matrix. For incoherent noise, asymptotically, $\Sigma \rightarrow \sigma^2 I$, yielding the following model for the observed Bartlett spectrum, dependent on the true parameter values:

$$\mathcal{F}(\theta : \theta_0) = S\mathcal{A}(\theta : \theta_0) + \sigma^2, \tag{1}$$

where we defined the generalized ambiguity

$$\mathcal{A}(\theta : \theta_0) = \left| h_\theta^H h_{\theta_0} \right|^2 .$$

Model (1) predicts an observed surface that is a scaled version of the ambiguity surface, plus an unknown constant noise term σ^2.

At each iteration, k, for each element of the population θ_n^k (i.e., for each set of estimates) we compute the generalized angle between the vector of the observed spectral samples,

$$\mathcal{F}^k = [\mathcal{F}(\theta_1^k) \ \ldots \ \mathcal{F}(\theta_N^k)]^T$$

and the ambiguity surface predicted for θ_n^k:

$$A(\theta_n^k) = [\mathcal{A}(\theta_1^k : \theta_n^k) \ \ldots \ \mathcal{A}(\theta_N^k : \theta_n^k)]^T .$$

The measure of correlation between the observed and predicted surfaces is based on equation (1), which can be written in vector form as

$$\mathcal{F}^k = SA(\theta_n^k) + \sigma^2 = \begin{bmatrix} A(\theta_n^k) & \mathbf{1} \end{bmatrix} \begin{bmatrix} S \\ \sigma^2 \end{bmatrix} .$$

Defining

$$M = \begin{bmatrix} A(\theta_n^k) & \mathbf{1} \end{bmatrix} ,$$

it is known that the best estimates of the unknown parameters are

$$\begin{bmatrix} \hat{S} \\ \hat{\sigma}^2 \end{bmatrix} = M(M^T M)^{-1} M^T S_k = \Pi_M S_k ,$$

where Π_M is the orthogonal projection matrix in the space spanned by the columns of M, yielding the following correlation between the corresponding estimated S_k and the observed vector:

$$C(\theta_k) = \frac{||\Pi_M \mathcal{F}^k||^2}{||\mathcal{F}^k||^2} .$$

This correlation measure has been used to evaluate each individual in each iteration of the genetic algorithm. Use of $C(\cdot)$ alone is not sufficient, since equally

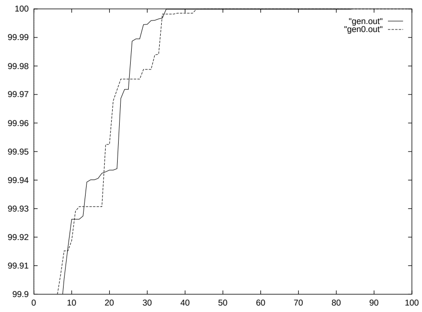

Figure 1: evolution of the Bartlett spectrum for the best individual in each iteration.

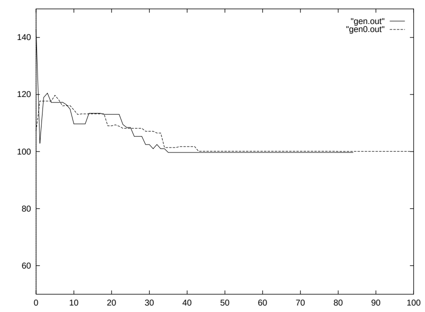

Figure 2: evolution of the estimated depth (indicated by the best individual in each iteration).

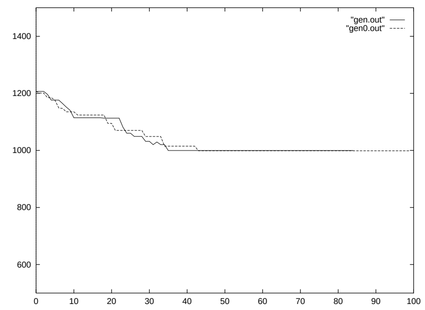

Figure 3: evolution of the estimated range (indicated by the best individual in each iteration).

high values (close to unit) can be reached either when θ_n^k is a good estimate, or when all the population is sampling the lower regions of the ambiguity surface. For this reason, the evaluation function combines the values of the Bartlett spectrum at each individual with the normalized correlation :

$$\mathcal{E}(\theta_n^k) = \mathcal{F}(\theta_n^k)C(\theta_n^k). \tag{2}$$

Using the product of the two indexes guarantees that high values of $\mathcal{E}$ must correspond both to a large value of $\mathcal{F}$ and to a good agreement of the sampled values to those predicted by the *a priori* model.

The function defined in (2) has been used to select the random set of individuals that are mutated and reproduced at each iteration, as well as the next generation. Selection of the next population is deterministic, chosing the set of N individuals that present the largest values of $\mathcal{E}$.

Figures 1–3 compare our algorithm with a standard implementation, which bases the selection of individuals (for mutation, reproduction and survival) only on the value of the Bartlett spectrum, revealing the faster convergence of the proposed technique. The number of individuals in each population is equal to 6. The actual source immersion and range are $R = 1000$ m and $D = 100$ m.

BIBLIOGRAPHY

[1] Thomas Bäck and Hans-Paul Schwefel. An overview of evolutionary algorithms for parameter optimization. *Evolutionary Comoutation*, 1, No. 1:1–23, 1993.

[2] Micahel Collins and W. A. Kuperman. Focalisation: environmental focusing and source localisation. *J. Acoust. Soc. Amer.*, 90:1410–1422, 1991.

[3] José M. F. Moura et M.João Rendas. Ambiguity function in sonar. In *Proceedings of the 1992 ASILOMAR Conference, Monterey, CA, USA*, Octobre 1992.

[4] Peter Gerstoff. Inversion of seismoacoustics data using genetic algorithms and a posteriori probability distributions. *J. Acoustic Soc. America*, 95:770–782, 1994.

[5] J. H. Holland. *Adaptation in natural and Artificial Systems*. AnnArbor: The University of Michigan Press, 1975.

[6] Georges Bienvenu José M. F. Moura, M. João Rendas. Environmental limits to source localization. In *Int. Conf. on Acoustic, Speech and Signal Processing95, Detroit, USA*, May 1995.

[7] M. João D. Rendas and José M. F. Moura. Ambiguity analysis in source localization with unknown signals. In *Int. Conf. on Acoustic, Speech and Signal Processing91, Toronto, Canada*, May 1991.

[8] J.F. Smith and S. Finette. Simulated annealing as a method of deconvolution for acoustic transients measured on a vertical array. *J. Acoustic Soc. America*, 94:2315–2325, 1993.

[9] H. Martynov Z.-H. Mochalopoulou and M. Porter. Simulated annealing and genetic algorithms for broadband source localization. In J. S. Papadikis, editor, *Proceedings of the 3rd European Conference on Underwater Acoustics*, pages 409–414. Found. for Research and Tech.–Hellas, Inst. of Appl. and Comp. Math., 1979.

[10] Anatoly A. Zhigljavsky. *Theory of Global Random Search*. Kluwer Academic Press, 1991.