OPTIMUM SUBARRAY CONFIGURATION USING GENETIC ALGORITHMS

J. Wang, H. Israelsson, R. G. North

Center for Monitoring Research, SAIC
1300 N. 17th Street, Suite 1450
Arlington, VA 22209, USA

ABSTRACT

Subarray configuration is not a trivial problem in array signal processing. A proper subarray configuration is important to improve the detectability of an array. A new searching algorithm, which is based on Genetic Algorithms (GA), for the optimum subarray configuration is proposed in this paper. Our preliminary application to a seismic array has indicated that the new algorithm can search a population of subarrays in a more efficient and robust way. The beamforming gain of the optimum subarray derived by GA is very close to the theoretical gain. Experimental results on signal detections have demonstrated that a beamforming recipe with optimum subarrays can provide further enhanced signal-to-noise ratio (SNR), compared to a recipe without subarray configuration. The approach proposed here can be easily extended to the weight determination problem for the weighted beamforming process by using multi-bit instead of 1-bit representation for each sensor in the chromosome model.

INTRODUCTION

Seismic arrays will play an important role in the International Monitoring System for the purpose of verifying compliance with the comprehensive nuclear test ban treaty of the United Nations. The ability of arrays to detect, locate, and characterize weak seismic signals has been well recognized. However, it has been a research topic to improve the detection and location capabilities of arrays.

Beamforming is a primary and powerful algorithm in array signal processing. The goal of beamforming is to increase SNR by suppressing incoherent background noise. The gain of beamforming is defined by:

$$G^2 = \frac{\sum_{i,j=1}^{N} w_i w_j C_{signal}(i,j)}{\sum_{i,j=1}^{N} w_i w_j C_{noise}(i,j)} \qquad (1)$$

where $C_{signal}(i,j)$ is the signal correlation between sensors i and j, $C_{noise}(i,j)$ is the corresponding noise correlation, w_i and w_j are sensor weights, and N is the number of sensors for beamforming.

In the common beamforming practice at seismic arrays, the weight of sensor is either 1 or 0. That is to say, using different combination of sensors, or called subarray configuration, is a conversional approach to improve the capability of detecting different signals. If signal waveforms are identical with proper time-shifting at N sensors and noise is spatially uncorrelated, the theoretical gain of beamforming would be:

$$G^2 = N \qquad (2)$$

In reality, however, the correlation of the signals recorded at different sensors are not perfect and may vary with sensor separation and signal frequency, also the noise may be partially or fully spatially correlated across the array [5]. The optimum beamforming gain could be achieved by subarray configuration which is based on the correlation characteristics of signal and noise at various frequency bands [4].

A new searching algorithm for subarray configuration is proposed in this paper. It is based upon a natural processing concept called genetic algorithm (GA) [3], which mimics the mechanics of natural selection and genetics. Experiment results using the proposed algorithm to a seismic array are presented for illustration.

METHODOLOGY

GA has been successfully applied in various artificial intelligence applications as a new optimization method [2]. Typically, GA consists of a bit-string representation of points called chromosomes in the search space, a misfit function to evaluate the search points, a set of operators for generating new chromosomes, and a stochastic assignment to control the genetic operators. To apply GA to subarray configuration, it is straightforward to represent a subarray with a binary coded chromosome of length M, which is the total number of sensors of array. In a chromosome, '1' means to turn on this sensor, '0' means not. For an M sensor array, the search space of subarray configuration will be 2^M. The goal of GA application to subarray configuration is to search a sensor grouping that will give an optimum gain of beamforming.

The proposed algorithm consists of the following steps:

1. Coherency Study -- Select a seismic signal and noise observed at all sensors of array, then calculate the correlation coefficient matrix between all sensor-pairs of the array in different frequency bands. For each seismic signal type, e.g., teleseismic P, a set of observations are selected for averaging purpose. The average coherence matrices of signal and noise are represented as $C_{signal}(i, j)$ and $C_{noise}(i, j)$.

2. Initialization -- Randomly generate an initial population of N_p chromosomes (subarrays) except two: one is full array, namely, all sensors are used for beamforming, and another is the subarray determined by average coherence distance. The average coherence distance is defined as the distance where the correlation coefficient drops to below 0.7 for seismic signals, and below 0.3 for noises.

3. Evaluation -- Evaluate the objective function of each chromosome within the population. The objective function in this approach is defined as the difference between the real gain computed from equation (1) and the expected theoretical gain given by equation (2):

$$dG = \frac{\sum_{i,j=1}^{N} C_{signal}(i,j)}{\sum_{i,j=1}^{N} C_{noise}(i,j)} - N \qquad (3)$$

where i, j are non-zero genes in the chromosome which are corresponding to the sensors used in the subarray.

4. Selection -- Select parent chromosomes based on the dG defined above for subsequent genetic operation (crossover) to produce pairs of child chromosomes. In the selection process, a parameter P_c is used to control the probability of a chromosome being selected.

5. Crossover -- Produce offspring through combining and exchanging genes of the parent chromosomes. The uniform crossover operation [6] is adopted here.

6. Mutation -- Randomly invert the bit of a gene on a chromosome with a mutation probability, P_m, to maintain diversity in the population.

7. Dynamic Population Control -- Replace some amount of weaker chromosomes within the population by new randomly generated ones. This operation is introduced in every K generations to avoid premature convergence, which usually leads to a local optimum [1].

8. Repeat Steps 3-7 until convergence of the objective function or a predefined number of generations has been reached.

APPLICATION TO A SEISMIC ARRAY

To test the proposed algorithm, a seismic array WRA, which was built in 1965 at Warramunga, Australia, has been selected as an experimental array. Figure 1 shows the geometry of the array WRA. It consists of 20 seismometers, arranged in two perpendicular lines, each of which has 10 sensors with a spacing of about 2.5 km.

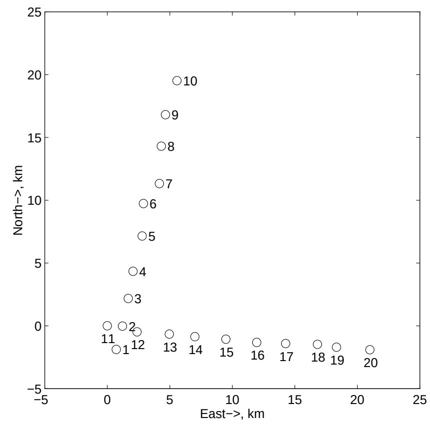

Figure 1. Geometry of the WRA array. Circles represent the location of sensors, and the nearby number is the sensor-id. The coordinates represent the distance to the reference sensor.

Coherency features of seismic signals and of noise (N) at WRA were firstly analyzed. For each of three types of seismic phases - teleseismic P (T), regional P (P) and regional S (S) - eight signals were selected for averaging purpose. Figure 2 gives an example of selected signals recorded at all sensors.

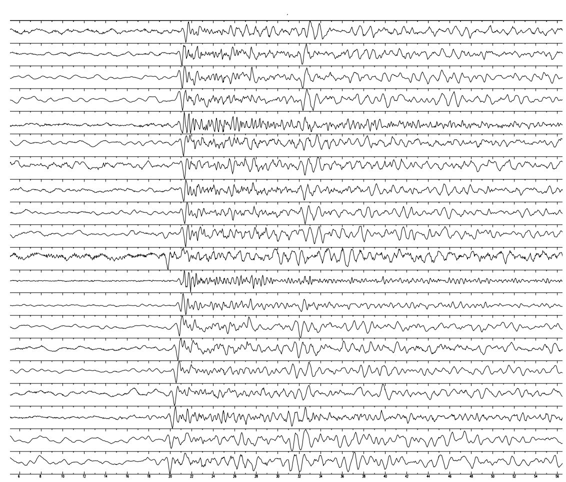

Figure 2. A sample of seismograms of a teleseismic P-phase recorded at all 20 seismometers at WRA. The waveform duration in this window is 26 seconds.

For each signal, waveforms of 20 channels were aligned with proper time-shifting and filtered in 8 frequency bands (0.5-1.0, 0.75-1.5, 1.0-2.0, 1.5-3.0, 2.0-4.0, 3.0-6.0, 4.0-8.0, 5.0-10.0 Hz). The correlation coefficient matrices between each sensor-pairs were calculated for different frequency bands. Figure 3 shows the average correlation coefficients for the teleseismic P phase as a function of separation distance, from which the coherence distance can be estimated.

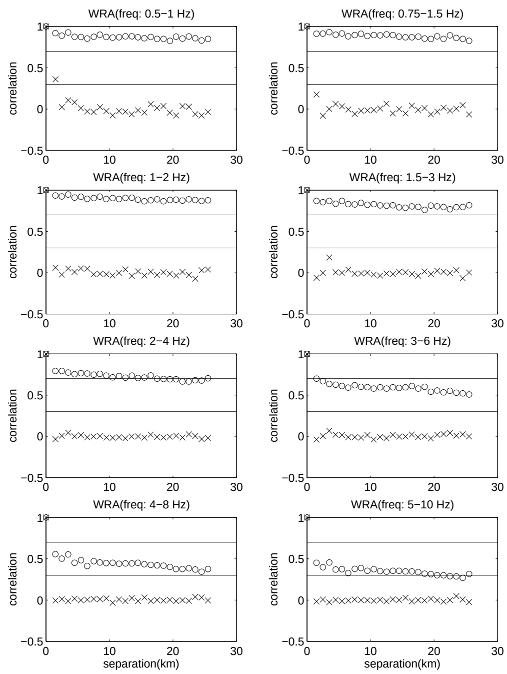

Figure 3. Average correlation coefficients of T type phase (open circles) and noise (crosses) as a function of sensor separation.

The average coherence distance of different signals in different frequency bands are listed in Table 1. From Table 1 we can see that the coherence distance decreases with increasing frequency. For signals in the two highest frequency bands, there is no coherence between any sensor-pairs at WRA. These two frequency bands should not be included in the beamforming process. Noise at low frequency band (0.5-1.0 Hz) is correlated within 2 km distance. To avoid coherent noise in the beamforming, that low frequency band should not be used either.

After calculating the average correlation matrices of different type signals in different frequency bands, the Genetic Algorithm was applied to search the optimum subarray configuration as described in last section. In this application, the chromosome length is 20 which corresponds to the number of sensors. The population size of chromosomes in each generation is 40. Among the 40 initial chromosomes, one represented full array (FA), one was derived from the average coherence distance (CD), and the rest were generated randomly. After up to 2500 generations of GA operation, the final optimum subarrays have been derived. Table 2 lists the final 8 subarrays in terms of bit-coded chromosomes. Table 3 lists the related beamforming gain relative to the theoretical gain. For comparison, the gain differences achieved by FA and CD configurations are also included in Table 3. Table 3 shows that the optimum subarrays determined by GA achieved the best performance relative to the theoretical gain.

Table 1: Average Coherence Distances of Different Phase Types at WRA

Frequency band (Hz)	T (km)	P (km)	S (km)	N (km)
0.50 - 1.0	26	26	4	2
0.75 - 1.5	26	26	3	0
1.0 - 2.0	26	26	2	0
1.5 - 3.0	26	15	0	0
2.0 - 4.0	18	3	0	0
3.0 - 6.0	2	0	0	0
4.0 - 8.0	0	0	0	0
5.0 - 10.0	0	0	0	0

Table 2: Subarray Configurations by GA

Phase type	Frequency band (Hz)	Subarray 00000000011111111112 12345678901234567890
T	0.75-1.5	11111111011111111111
	1.0 - 2.0	11111111111111111111
	1.5 - 3.0	11110111111111111111
	2.0 - 4.0	11011110010101110010
P	0.75-1.5	00110111111111011010
	1.0 - 2.0	00111111111110111010
	1.5 - 3.0	00111111100100011110
S	0.75-1.5	00101001001101010000

Table 3: Comparison of Differences between Real Gain and Theoretical Gain for Different Configurations

Phase type	Frequency bands (Hz)	dG		
		FA	CD	GA
T	0.75-1.5	1.49	1.49	1.54
	1.0 - 2.0	-0.66	-0.66	-0.66
	1.5 - 3.0	-2.51	-1.34	-0.80
	2.0 - 4.0	-3.81	-1.45	-0.33
P	0.75-1.5	-6.70	-6.70	-0.90
	1.0 - 2.0	-5.23	-4.60	-0.40
	1.5 - 3.0	-7.53	-5.6	-0.90
S	0.75-1.5	-10.17	-2.1	-0.80

Depending on the goal of array operation, different subarray should be used for beamforming to detect different signals of interests. To evaluate the detection performance of the optimum subarrays derived by GA, two control files (‘recipe’) for beamforming detection were constructed with a systematic approach [7]. One recipe employs all optimum subarrays derived by GA and another employs full array (FA). All detection beams are steered with an omnidirectional pattern and 3dB resolution in the wavenumber plane.

In the experiments for testing these two recipes, simulated on-line detection operation was conducted with one-day continuous waveform data. Testing results show that the average SNR of the detected real signals, which were confirmed by analysts, is 44.3 by using the GA recipe, while 38.4 by using the FA recipe. At same time, however, the average SNR of false detections (noise) is 6.95 for the GA recipe, while 7.60 for the FA recipe. This means that the beamforming recipe with the optimum subarrays can improve the SNR of signals while depressing noise amplitudes more efficiently than the recipe with the full array. The improvement of SNR by GA recipe relative to the FA recipe can be obtained from the ratio (44.3/6.95)/(38.4/7.60), which equals 1.26 times.

SUMMARY

Genetic Algorithms can be applied to the optimum subarray configuration. Our preliminary work indicated that the proposed algorithm can search a population of subarrays in a more efficient and robust way. The performance of beamforming gain is very close to the theoretical gain. Experimental results on signal detections have demonstrated that a beamforming recipe with optimum subarray configuration can provide further enhanced SNR compared to a recipe without subarray configuration.

This approach can be easily extended to the optimum weight determination problem for the weighted beamforming. Instead of using 1-bit in chromosome to represent each sensor, a multi-bit representation should be used. For example, if we want to choose 8 different levels of weight between 0 and 1 for the weighted beamforming at WRA, we can use 3-bit length for each sensor, and the total length of chromosome will be 60 bits. The remaining operations for searching optimum weights will be same.

REFERENCES

[1] Chow K. H. and Liou M. L. “Genetic motion search algorithm for video compression”. *IEEE Trans. Circuits Syst. Viedo Technol.*, vol. 3, no. 6, pp. 440-445, Dec. 1993.

[2] Davis L., Editor. *Genetic Algorithms and Simulated Annealing, Research Notes in Artificial Intelligence*. Pitman, London, 1990.

[3] Holland J. H. *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, 1975.

[4] Kvaena T. “On exploitation of small-aperture NORESS type arrays for enhanced P-wave detectability”. *Bull. Seism. Soc. Am.* **79**, 888-900, 1989.

[5] Mykkeltveit S., Astebϕl K., Doornbos, D. J., and Husebye E. S. “Seismic array configuration optimization”. *Bull. Seism. Soc. Am.* **73**, 173-186, 1983.

[6] Syswerda G. “Uniform crossover in genetic algorithms”. *Proc. 3rd Int. conf. genetic Algorithms and Applications,* CA, pp. 2-9, 1989.

[7] Wang J., Israelsson H., and Carter J. “A systematic approach to designing detectors for diverse seismic arrays”. *Seismological Research Letters*, **67**, 60, 1996.