Full waveform microseismic inversion using differential evolution algorithm

Lijun Zhu, Entao Liu, and James H. McClellan
Center for Energy & Geo Processing (CeGP) at Georgia Tech and KFUPM
Atlanta, GA, USA
lijun.zhu@gatech.edu, {entao.liu, jim.mcclellan}@ece.gatech.edu

***Abstract*—An accurate and fast estimation of microseismic activities from passive microseismic data is a crucial issue to many oil and gas applications. Traditional methods are mainly based on manual picking of the arrival times which require a relative high signal-to-noise ratio (SNR) to produce reliable results. When the sensor array is deployed on the surface, the microseismic events have low magnitude and might be buried in strong ambient noise. A compressive sensing scheme has been introduced to implement seismic source parameters estimation, including location (i.e., hypocenter) and moment tensor (MT). Although this scheme is efficient and accurate, it entails computing the whole dictionary composed of Green's functions in advance, which brings a high computational overhead and prohibitive storage burden. In this work, we propose the differential evolution (DE) algorithm for solving the inverse problem on the fly which avoids generating and storing the whole dictionary.**

I. Introduction

Microseismic monitoring has become a primary tool for many applications in developing oil and gas reservoirs, such as hydraulic fracturing, water injection, and well completion operations. Seismic parameters, such as location, origin time, and the source mechanism (represented by MT), provide valuable information to improve the aforementioned operations, and eventually increase the productivity of the reservoir[1]. The process of microseismic monitoring implemented by passive and continuous recording using a surface geophone array is becoming increasingly popular due to its relative low cost and wide aperture compared with downhole sensor arrays.

Traditional techniques for locating microseismic events based on picking the first arrival times require good SNR. However, the microseismic data collected by a surface array are typically rather noisy so that even manual picking becomes inaccurate or fails. Owing to recent advances in computation power, full waveform matching of both P-waves and S-waves has now become feasible. By minimizing the least squares mismatch between the measured data and the predicted data, we can estimate the hypocenter and origin time of a microseismic event. The Green's functions with respect to a location provide an impulse response on the geophone array. Throughout this paper, we abuse the concept of a Green's function slightly to obtain the array sensor responses as the convolution with an assumed source time function. Moreover, the MT of microseismic events contains valuable information about the reservoir and the fracturing mechanisms, which is difficult to retrieve by conventional schemes[2]. The advantage of full waveform matching based methods is their ability to recover the MT simultaneously with little extra effort.

As to the optimization, a commonplace gradient-based method is not applicable in this scenario for the following reasons: (1) It is almost impossible to analytically compute the gradient of the objective function, especially for a non-homogeneous media. (2) The numerical expression of the gradient needs an explicit form for the whole dictionary [3]. For large scale problems, the gradient is extremely computationally expensive. Hence, an efficient gradient-free algorithm for this global optimization is highly desirable. Currently, an exhaustive search over all grid nodes, grid search (GS), is conducted to estimate the hypocenter and MT. This requires computing the Green's functions of each node on the fly or storing a huge pre-computed dictionary on hard drives. In order to control the size of the dictionary and the computational burden, people usually carry out this point-by-point search on a limited region, or a coarse grid. It is known that the coherence of the Green's functions, considered as columns in the dictionary, is dependent on the events' spatial locations. Concretely, columns whose spatial locations are close to each other are highly coherent. If the Green's function associated with a specific node match the data poorly, the nearby nodes are very unlikely to match the data well. This fact suggests that an exhaustive search is not necessary for this highly redundant dictionary. The DE algorithm is adopted here to recover the microseismic parameters. Actually, the DE algorithm has already attracted attention from the geophysical community for its robustness and ease of implementation [4], [5]. Instead of evaluating the cost function on all nodes, the DE method only evaluates on a number of candidate nodes which are small fraction of all the grid nodes in the search space. In this manner, the overall computational cost of searching (including generating the Green's functions) is significantly decreased. Thus, an accurate and efficient search over a larger area, or in a finer resolution, becomes feasible.

II. Dictionary based on Green's functions

Besides the geophysical model, the received microseismic signals on a sensor are dependent on the seismic source parameters (e.g., location, origin time, and moment tensor) via elastic wave equations. In order to formulate this parameter estimation problem as

978-1-4799-7591-4/15/$31.00 ©2015 IEEE

a linear inverse problem, we adopt the dictionary matrix structure of $\mathbf{G}$, where each column is a Green's function corresponding to a certain node and one of the free parameters in the MT matrix [6]. By [7] the displacement field observed at location $\mathbf{x}$ due to a source at ξ with source mechanism given by the scalar MT matrix M_{pq} can be written as

$$\begin{aligned} u_n(\mathbf{x},t) &= \sum_{p=1}^{3}\sum_{q=1}^{3} M_{pq}\left[s(t) * \frac{\partial}{\partial \xi_q} G_{np}(\mathbf{x},t;\xi,t_0)\right] \\ &= \sum_{p=1}^{3}\sum_{q=1}^{3} M_{pq} g_{npq}(\mathbf{x},t;\xi,t_0), \end{aligned} \tag{1}$$

where $s(t)$ is the source time function and G_{np} the Green's function. The index n refers to the receiver component and t_0 is the origin time of the source. We can express (1) as matrix multiplication

$$\mathbf{u} = \mathbf{G}\mathbf{m}, \tag{2}$$

where $\mathbf{m}$ is the vector of coefficients. The MT matrix M is a symmetric 3×3 matrix. For each node (a potential hypocenter of the microseismic event), there are six columns according to the six free parameters of M. In total, the number of columns in $\mathbf{G}$ is $6 \times N$, where N is the number of nodes in the region of interest. For convenience, the columns of $\mathbf{G}$ can be denoted as $[\mathbf{G}[1], \ldots, \mathbf{G}[N]]$, where each submatrix $\mathbf{G}[i]$ comprises 6 columns corresponding to the i^{th} node. The displacement field on the array $\mathbf{u}$ is of size $N_t \times N_r \times 3$, where N_t is the number of samples in a single trace and N_r is number of 3-component (3-C) sensors in the array. Empirically, this linear inverse problem is highly sparse, i.e., there are only a few events in the time interval we are examining.

III. Searching algorithm

In order to invert the source parameters, i.e., the vector $\mathbf{m}$, block sparse inverse problem solving methods such as GS are often employed. The major shortcoming of this approach is that the whole dictionary $\mathbf{G}$ must be computed and stored. This issue can be alleviated by reducing the number of grid nodes, i.e., using a coarse grid or a limited monitoring region. This approach is clearly a trade-off that may cause inaccuracy and missed detections for many off-grid events due to the coarse grid. Moreover, when the examining time window contains more than one event, the GS can only pick one event that contributes the most significant displacement field. To address these issues, we adapt the differential evolution (DE) algorithm [8] to solve this inverse problem.

In the literature, the use of evolutionary algorithms in geophysical applications is not new. For example, the particle swarm method [9] and the DE algorithm [4], [5] have been applied to solve the microearthquake localization with a semblance based cost function. Simulation shows the gradient-free DE algorithm is robust and effective for cumbersome global optimization, which is carried out as follows.

Initialization: We randomly pick a initial population of agents consisting of a set of parameter vectors in the appropriate ranges. As a rule of thumb, we draw all components in the parameter vector from a uniform distribution, when we have no prior knowledge of the parameters. The population size (denoted by P) is normally taken as 5 to 10 times D [8], which is the size of a parameter vector, and remains unchanged.

Mutation: The DE generates new parameter vectors by adding the weighted difference between two population vectors to a third vector. Three distinct vectors are randomly chosen from the existing population, and a mutant vector is computed perturbing one vector with a scaled difference of the other two:

$$\mathbf{v}_p = \mathbf{x}_{p_1} + F(\mathbf{x}_{p_2} - \mathbf{x}_{p_3}), \tag{3}$$

where $\mathbf{v}_p$ represents the p^{th} mutant vector; and $\mathbf{x}_{p_1}$, $\mathbf{x}_{p_2}$, and $\mathbf{x}_{p_3}$ represent the three distinct randomly chosen parameter vectors. The scaling factor $F \in [0, 2]$ denotes the mutation factor, which controls the amplification of the differential variation $(\mathbf{x}_{p_2} - \mathbf{x}_{p_3})$. [5]

Crossover: This step picks components of the competing vector in the current vector and the mutant vector to increase the diversity. Let p_j be the j^{th} realization of a random variable of a given distribution. For $j = 1, \ldots, D$ the parameter vector u is given by

$$u_j = \begin{cases} v_j & \text{if } p_j \le C \text{ or } j = RI \\ x_j & \text{otherwise} \end{cases} \tag{4}$$

where u_j, v_j, and x_j represent the j^{th} element of a parameter vector in the competing population, mutant population, and current population, respectively. C is the crossover constant $\in [0, 1]$. RI is a randomly chosen index $\in \{1, 2, \ldots, D\}$ which ensures the competing u_j gets at least one parameter from the mutant vector.

Selection: Decide whether or not the competing vector should become a member of the next generation. The competing vector $\mathbf{u}_i$ is compared to the current vector $\mathbf{x}_i$ (for $i = 1, \ldots, P$) by evaluation of the cost function. A one-by-one comparison of the competing population and the current population is carried out for all vectors. The parameter vectors with a lower value of the cost function compose the next generation.

Practically, the population of agents can cover the true hypocenter in a small number of iterations. When the true hypocenter is contained in the population, we can use a combination of the DE algorithm and exhaustive search within the small bounded region, rather than continue running the DE for too many iterations. Assume the origin time is known, the cost function here is only determined by node location. For a presumed l^{th} node, we first compute the least squares solution of MT by

$$\hat{\mathbf{m}}(l) = (\mathbf{G}^H[l]\mathbf{G}[l])\mathbf{G}^H[l]\mathbf{u}. \tag{5}$$

Then the cost function is

$$J(l) = \|\mathbf{u} - \mathbf{G}[l]\hat{\mathbf{m}}(l)\|. \tag{6}$$

IV. Simulation and Validation

A. Simulation setup

For simplicity, we only test the proposed method in a known homogeneous elastic model. In general, if an accurate geophysical model is given, the proposed method is expected to have similar performance. An uniform 15×15 square array of 3-C geophones is deployed on the surface as in Fig. 1 with 40 m spacing. The source signal is modelled by a Ricker wavelet with center frequency at 40 Hz. The synthetic data are generated using the far field formulation in [10] plus white Gaussian noise. The monitored region of interest is underneath the array, which is a cube that covers -300 m to 900 m in both x and y directions, and 400 m to 1000 m in depth. The true location of the event is set to be $(280, 280, 800)$ m.

The noise model used here is additive Gaussian white noise. Here we define a strict peak signal-to-noise ratio (PSNR) as the ratio between the maximum value on all traces ($D_{\max}$) over the standard deviation of the white Gaussian noise (σ) as

$$\text{PSNR} = 20 \log_{10} \frac{D_{\max}}{\sigma}. \tag{7}$$

A noisy seismic gather of 25 dB PSNR is shown in Fig. 2. Note that the z component in Fig. 2(c) seems to be more noisy than the other two, because the magnitude of z component is the smallest among the three. The relative signal amplitudes among the three axes are determined by the MT of the source. For the common double couple source mechanism, the MT is dependent on three parameters—the slip, dip, and rake (see details in [7]). Generically, we assume the source is a double couple whose slip, dip, and rake are 30°, 30°, and 180°, respectively. Then we calculate its MT

$$\text{MT} = \begin{bmatrix} 0.4330 & -0.2500 & 0.7500 \\ -0.2500 & -0.4330 & 0.4330 \\ 0.7500 & 0.4330 & 0.0000 \end{bmatrix}.$$

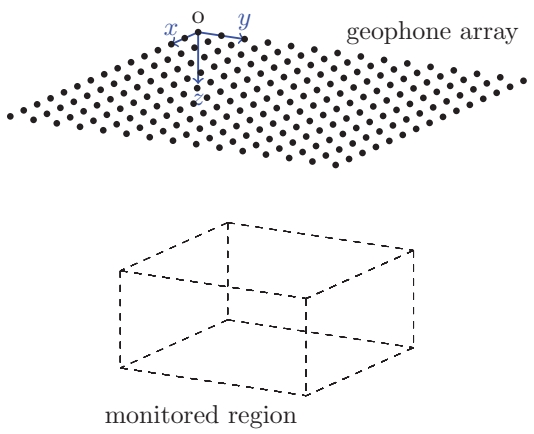

Fig. 1. Layout of the testing array of 15×15 surface geophones with 40 m spacing.

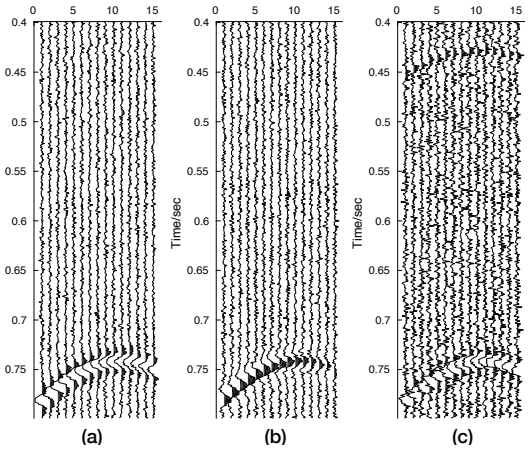

Fig. 2. Noisy signals (PSNR = 25 dB) on a line of 15 geophones: x, y, and z components.

In this study, we denote the computation needed for generating Green's functions of one node and implementing MT inversion according to this node as one computation unit. If the GS is applied for source location estimation with 40 m grid spacing, the total is $(\frac{1200}{40})(\frac{1200}{40})(\frac{600}{40}) = 13500$ computation units are needed to perform exhaustive search. For finer spacing or a larger monitoring area, the GS method can quickly become unfeasible. On the other hand, the DE method requires a population of agents, whose cardinality is five to ten times the dimension of the search space [8] which is only three (x, y, and z) in this case. In each iteration, it computes the Green's functions for each agent and conducts MT inversion with the observed data. The DE method used here can be modified so that any off-grid point in a population will be projected onto the nearest grid node. As a result, if the true hypocenter is on a off-grid point, the method will only find the nearest grid node by this projection. This technique effectively reduces the size of search space from a continuous region into a discrete grid.

B. Convergence rate

To understand the convergence rate of the pure DE method for this search problem, a noiseless simulation is conducted as shown in Fig. 3 with 122 Monte Carlo experiments. After 100 iterations with a population of 30 agents in each, the mean error of hypocenter estimation is 9.978 m with a standard deviation of 21.86 m. Thus, we have approximately 95% confidence that the optimal solution of the DE method will be within 40 m (one grid spacing) of the true location. In other words, if 5% uncertainty is acceptable in the hypocenter estimation, the DE method only requires $30 \times 100 = 3000$ computation units instead of 13500 in GS to locate the microseismic event. Theoretically, the convergence rate of DE is not guaranteed and usually slows down after a certain number of iterations. Taking Fig. 4 as an example, the DE method reaches the true location at iteration 60, but the population of agents does not converge to the true location until iteration 165. Thus, it

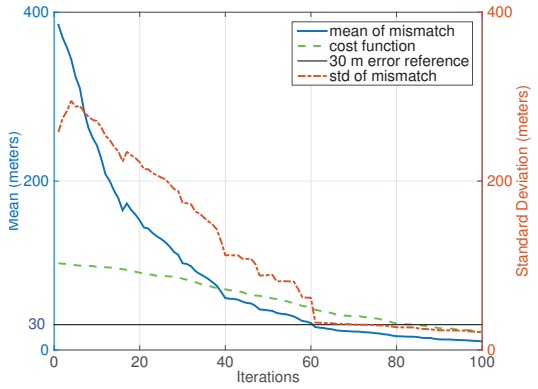

Fig. 3. Error between estimated and true hypocenters versus the number of iterations. The trend of the cost function (dashed line) is plotted as reference.

is not sensible to run the DE for too many iterations, since a much smaller number of iterations already provides an estimation with an acceptably small deviation.

The mean and standard deviation is 27.59 and 32.16 m after 60 iterations in Fig. 3. Therefore, we can use GS to search a $5 \times 5 \times 5$ cube after 60 iterations of DE. In this case, only $60 \times 30 + 5^3 = 1925$ computation units are needed for DE + GS method to find the correct event location with confidence greater than 95% .

C. Performance against noise

The performance of the DE method is studied under different noise levels. Using a Monte Carlo method of 12 repeated experiments, the DE method of the same setup is tested for PSNRs from 0 dB to 40 dB. The performance shown in Fig. 5 is measured as the standard deviation of the error between the estimated location and the true location for 12 experiments. As expected, the lower the PSNR level, the higher the standard deviation of the error will be. The standard deviation of the error of all tested PSNR cases is

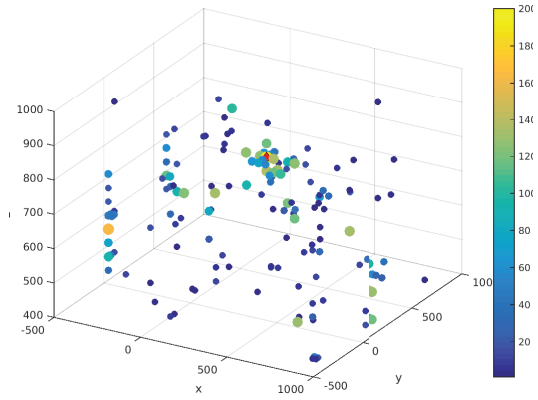

Fig. 4. Distribution of DE agents in 3D with respect to number of iterations for 200 iterations. A larger ball size indicates more iterations. The red diamond indicates the true hypocenter.

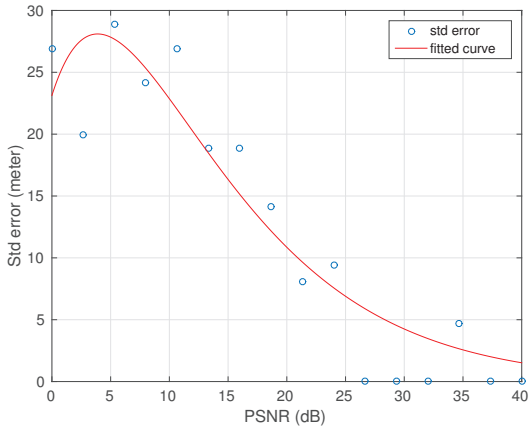

Fig. 5. Standard deviation of the location error using the DE method for different PSNR levels.

within one grid spacing. This shows that this DE method of full waveform matching is rather robust under noise.

We note that the good performance demonstrated in Fig. 5 is obtained under the assumption that the origin time of the event has been correctly picked. In a low PSNR scenario, this is nearly impossible. The DE method can be modified to include the detection of origin time by simply adding one more temporal parameter in the spatial search domain. However, adding one more dimension in time will greatly increase the already huge size of dictionary. The quantitative study of time origin inversion is beyond the scope of this paper.

V. Conclusion

Locating microseismic events and inverting the MT from the surface data is important for many seismic applications. In this study, we show that using differential evolution (DE) one can find an accurate hypocenter and MT with significantly less computation than with the grid search (GS) method. Additionally, the DE followed by GS in a small region can improve the computational efficiency further. In a case study, we show that the DE method converges in tens of iterations and is robust over a range of PSNR levels.

Acknowledgement

This work is supported by the Center for Energy and Geo Processing (CeGP) at Georgia Tech and by King Fahd University of Petroleum and Minerals (KFUPM).

REFERENCES

[1] P. M. Duncan and L. Eisner, "Reservoir characterization using surface microseismic monitoring," *Geophysics*, vol. 75, no. 5, pp. 75A139–75A146, 2010.

[2] B. D. E. Dando, K. Chambers, and R. Velasco, "A robust method for determining moment tensors from surface microseismic data," in *SEG Technical Program Expanded Abstracts 2014*, 2014, pp. 2261–2266.

[3] I. V. Rodriguez, M. D. Sacchi, and Y. J. Gu, "Continuous hypocenter and source mechanism inversion via a Green's function-based matching pursuit algorithm," *The Leading Edge*, pp. 334–337, March 2010.

[4] K. Chambers, J. Kendall, S. Brandsberg-Dahl, and J. Rueda, "Testing the ability of surface arrays to monitor microseismic activity," *Geophysical Prospecting*, vol. 58, no. 5, pp. 821–830, 2010.

[5] H. N. Gharti, V. Oye, M. Roth, and D. Kühn, "Automated microearthquake location using envelope stacking and robust global optimization," *Geophysics*, vol. 75, no. 4, pp. MA27–MA46, 2010.

[6] I. V. Rodriguez and M. D. Sacchi, "Microseismic source imaging in a compressed domain," *Geophysical Journal International*, vol. 198, pp. 1186–1198, 2014.

[7] K. Aki and P. G. Richards, *Quantitative Seismology*, 2nd ed. Sausalito CA: University Science Books, 2009.

[8] R. Storn and K. Price, "Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 341-359, 1997.

[9] S. R. Lagos, J. I. Sabbione, and D. R. Velis, "Very fast simulated annealing and particle swarm optimization for microseismic event location," in *SEG Denver 2014 Annual Meeting*, 2014, pp. 2188–2192.

[10] C. H. Chapman, *Fundamentals of Seismic wave propagation*. Cambridge, UK: Cambridge University Press, 2004.