Pattern Search Algorithms for Surface Wave Analysis

Xianhai Song
Changjiang River Scientific Research Institute, Wuhan, Hubei 430010
China

1. Introduction

In recent years Rayleigh waves have been used increasingly as a nondestructive tool to obtain near-surface S-wave velocity, one of the key parameters in environmental and engineering applications (Miller et al., 1999a, 1999b; Park et al., 1998, 1999; Xia et al., 1999, 2003; Foti et al., 2002; Beaty and Schmitt, 2003; Tian et al., 2003; Zhang and Chan, 2003; Lin and Chang, 2004; O’Neill et al., 2003; Ryden et al., 2004; and Song et al., 2006). Utilization of surface wave dispersive properties may be roughly divided into three steps: field data acquisition, dispersion-curve picking, and inversion of phase velocities (Xia et al., 2004). Inversion of Rayleigh wave dispersion curves is one of the key steps in surface wave analysis to obtain a shallow subsurface S-wave velocity profile (Rix, 1988; Lai et al., 2002). However, inversion of high-frequency Rayleigh wave dispersion curves, as with most other geophysical optimization problems, is typically a nonlinear, multiparameter inversion problem. Dal Moro et al. (2007) have demonstrated the high nonlinearity and multimodality by a synthetic model. Consequently, local optimization methods, e.g. matrix inversion, steepest descent, conjugate gradients, are prone to being trapped by local minima, and their success depends heavily on the choice of a good starting model (Boschetti et al., 1996) and the accuracy of the partial derivatives. Thus, global optimization methods that can overcome this limitation are particularly attractive for surface wave analysis (e.g., Meier and Rix, 1993; Yamanaka and Ishida, 1996; and Beaty et al., 2002).

Pattern search algorithms are a direct search method well suitable for the global optimization of highly nonlinear, multimodal objective functions (Lewis and Torczon, 1999). Pattern search algorithms have recently been used and tested to optimize complex mathematical problems characterized by the large numbers of local minima and/or maxima (Lewis and Torczon, 2000). However, few attempts have been made to address real-world geophysical problems, especially for the inversion of surface wave data. In this chapter, I first implemented and tested a Rayleigh wave dispersion curve inversion scheme based on GPS Positive Basis 2N, a commonly used pattern search algorithm. Incorporating complete poll and complete search strategies based on GPS Positive Basis 2N into the inverse procedure greatly enhances the performance of pattern search algorithms because the two steps can effectively locate the promising areas in the solution space containing the global minima and significantly reduce the computation cost, respectively. The calculation efficiency and stability of the inversion scheme are tested on three synthetic models. Secondly, effects of the number of data points, the reduction of the frequency range...
of the considered dispersion curve, errors in P-wave velocities and density, the initial S-wave velocity profile as well as the number of layers and their thicknesses on inversion results are investigated to further evaluate the performance of the proposed approach. Thirdly, I implemented an investigation to fully exploit and utilize the potentiality of pattern search algorithms and to further enhance their performance for surface wave analysis. I investigate effects of different inversion strategies, initial mesh size and final mesh size, expansion factor and contraction factor, as well as inclusion of noise in surface wave data on the performance of the approaches, by three synthetic earth models. Fourthly, a comparative analysis with genetic algorithms is made to further highlight the advantages of the proposed inverse procedure. Finally, the performance of pattern search algorithms is verified by a real-world example from Henan, China.

2. Fundamentals on pattern search algorithms

As described above, pattern search algorithms are a direct search method well capable of solving global optimization problems of irregular, multimodal objective functions, without the need of calculating any gradient or curvature information, especially for addressing problems for which the objective functions are not differentiable, stochastic, or even discontinuous (Torczon, 1997). As opposed to more traditional local optimization methods that use information about the gradient or partial derivatives to search for an optimal solution, pattern search algorithms compute a sequence of points that get closer and closer to the globally optimal point. At each iteration, the algorithms poll a set of points, called a mesh, around the current point — the point computed at the previous iteration of the algorithms, looking for a point whose objective function value is lower than the value at the current point. If this occurs, the poll is called successful and the point they find becomes the current point at the next iteration. If the algorithms fail to find a point that improves the objective function, the poll is called unsuccessful and the current point stays the same at the next iteration. The mesh is formed by adding the current point to a scalar multiple (called mesh size) of a set of vectors (called a pattern). In addition to polling the mesh points, pattern search algorithms can perform an optional step at every iteration, called search. At each iteration, the search step applies another optimization method to the current point. If this search does not improve the current point, the poll step is performed (Lewis and Torczon, 2002).

Pattern search algorithms use the Augmented Lagrangian Pattern Search (ALPS) to solve nonlinearly constrained problems (Conn et al., 1991). The ALPS attempts to solve a nonlinear optimization problem with nonlinear constraints, linear constraints, and bounds using Lagrange multiplier estimates and penalty parameters (Lewis and Torczon, 2002). ALPS begins by using an initial penalty parameter and a starting point X0 for starting the optimization process. Pattern search algorithms minimize a sequence of subproblems, which are approximations of the original problem. When the subproblems are minimized to a required accuracy and satisfy feasibility conditions, the Lagrangian estimates are updated. Otherwise, the penalty parameter is increased by a penalty factor. This results in a new subproblem formulation and a minimization problem. These steps are repeated until the stopping criteria are met. For a more detailed description of the algorithms, the interested reader can refer to several excellent publications that extensively cover the subject (e.g., Conn et al., 1991; Torczon, 1997; Lewis and Torczon, 1999, 2000, 2002; Audet and Dennis, Jr., 2003; and Kolda et al., 2003).
3. Pattern search algorithms for surface wave analysis

We implemented a series of MATLAB tools based on the GADS (Genetic Algorithms and Direct Search) Optimization Toolbox of MATLAB 7.1 for surface wave analysis. We have worked on SWIPSA, a software package for Surface Wave Inversion via Pattern Search Algorithms. The codes being developed aim at performing nonlinear inversion of the fundamental and/or higher mode Rayleigh waves using pattern search algorithms.

In the current study, we implemented and tested a Rayleigh wave dispersion curve inversion scheme based on GPS Positive basis 2N, a commonly used pattern search algorithm. GPS Positive basis 2N is the Generalized Pattern Search (GPS) algorithm with the maximal positive basis set 2N vectors, where N is the number of independent variables (which in our case involve S-wave velocities) in the optimization problem. The algorithm uses fixed direction vectors to compute the set of points forming the mesh.

Complete poll method based on GPS Positive basis 2N is incorporated into the inversion procedure to discover the most promising domain in the solution space for a good valley. In general, if a pattern search algorithm finds a point in the mesh that improves the objective function at the current point, it stops the poll and sets that point as the current point for the next iteration. When this occurs, some mesh points might not get polled. Some of these unpollled points might have an objective function value that is even lower than the first one the pattern search finds. Especially for problems in which there are the large numbers of local minima and/or maxima of the object functions, it is sometimes preferable to make the pattern search poll all the mesh points at each iteration and choose the one with the best objective function value (Lewis and Torczon, 1999). Although a complete poll can make the pattern search run significantly longer, this enables the algorithm to explore more points at each iteration and thereby potentially avoid a local minimum that is not the global minimum.

Complete search method based on GPS Positive basis 2N is also adopted to reduce the total function evaluations (the number of times the objective function was evaluated) and the number of iterations (Lewis and Torczon, 1999). This choice is justified by the fact that computation time and efficiency are improved in our latter tests.

The expansion factor and contraction factor control how much the mesh size is expanded or contracted at each iteration. During the inversion procedure, we adopt the values suggested by the GADS Toolbox. We set the initial mesh size at 1, expansion factor at 2, contraction factor at 0.5, initial penalty at 10, and penalty factor at 100, which means the GPS algorithm multiplies the mesh size by 2 after each successful poll and multiplies the mesh size by 0.5 after each unsuccessful poll.

In this chapter, we focus our attention on inversion results of fundamental-mode Rayleigh wave dispersion curves for a near-surface S-wave velocity profile by fixing other parameters, namely P-wave velocities (or Poisson’s ratio), density, and thickness, to their known values or good estimates. The reduction of the computation time and the fact that the S-wave velocity is the most important parameter affecting the Rayleigh wave propagation (Xia et al., 1999) are the main reasons for these choices.

To simulate more realistic cases where no a priori information is available, an initial S-wave velocity profile (a starting point X0) with a uniform constant-velocity half-space is adopted. This is a simple and efficient scheme obtained by averaging maximum and minimum phase velocities, which is given by Eq. (1):

\[ V_{S0} = \left( \frac{V_{R}^{\text{max}} + V_{R}^{\text{min}}}{2} \right) \quad \text{(For all layers)} \]
The pattern search algorithm presented here is designed to treat the explicit lower and upper bound constraints that are present in the least-squares problem. To fully evaluate the capability of the algorithm and further simulate realistic cases where no a priori information is available, we use a wider search scope of the solution space. The lower and upper bounds of the search areas are 100 and 1000 m/s, respectively, for every layer in all of the latter tests. By design, these boundary values are approximately 100-150% off the true values.

The procedure will set out to find the global minimum of rms (root-mean-square) error misfit between the measured and the predicted phase velocities. The objective function is defined as:

\[ F = \frac{\| \mathbf{V}_R^{\text{obs}} - \mathbf{V}_R^{\text{theo}} \|}{\sqrt{m}} \]  

where \( \mathbf{V}_R^{\text{obs}} \) is an \( m \times 1 \) vector of the observed Rayleigh wave phase velocities, \( \mathbf{V}_R^{\text{theo}} \) is an \( m \times 1 \) vector of the theoretical Rayleigh wave phase velocities, \( m \) is the number of dispersion points (phase velocities versus frequency), and \( \| \cdot \| \) denotes the Euclidean norm of a vector. Forward modeling of the dispersion curves is based on Knopoff’s method (Schwab and Knopoff, 1972).

The algorithm is terminated at the 300th iteration in our inversion procedure or when the error misfit reaches a certain previously fixed value.

### 4. Synthetic data inversion

To examine and evaluate the efficiency and stability of the pattern search algorithm, three synthetic earth models are used. These models are designed to simulate situations commonly encountered in shallow engineering site investigations. As shown in Table 1, Model A, which consists of one homogeneous layer lying over a half-space with downwardly increasing S-wave velocities, represents a simple two-layer geologic structure. Model B (Table 2) with three layers over a half-space represents a simple multilayer ground condition. Model C (Table 3), characterized by four layers lying over a half-space with a 2-m-thick low velocity layer between two higher S-wave velocity layers, models a real complex subsurface structure actually analyzed and validated on a roadbed test site in Henan, China (Song et al., 2006).

<table>
<thead>
<tr>
<th>Layer Number</th>
<th>S-wave velocity (m/s)</th>
<th>P-wave velocity (m/s)</th>
<th>Density (g/cm³)</th>
<th>Thickness (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>416</td>
<td>1.8</td>
<td>6</td>
</tr>
<tr>
<td>Half-space</td>
<td>500</td>
<td>1041</td>
<td>2.0</td>
<td>Infinite</td>
</tr>
</tbody>
</table>

Table 1. Model A. Earth model parameters of one layer over a half-space.

<table>
<thead>
<tr>
<th>Layer Number</th>
<th>S-wave velocity (m/s)</th>
<th>P-wave velocity (m/s)</th>
<th>Density (g/cm³)</th>
<th>Thickness (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>416</td>
<td>1.8</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>250</td>
<td>520</td>
<td>1.8</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>300</td>
<td>625</td>
<td>1.8</td>
<td>6</td>
</tr>
<tr>
<td>Half-space</td>
<td>400</td>
<td>833</td>
<td>1.8</td>
<td>Infinite</td>
</tr>
</tbody>
</table>

Table 2. Model B. Earth model parameters of three layers over a half-space.
Table 3. Model C. Earth model parameters of four layers over a half-space.

<table>
<thead>
<tr>
<th>Layer Number</th>
<th>S-wave velocity (m/s)</th>
<th>P-wave velocity (m/s)</th>
<th>Density (g/cm³)</th>
<th>Thickness (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>346</td>
<td>1.8</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>150</td>
<td>765</td>
<td>1.8</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>663</td>
<td>1.8</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>300</td>
<td>995</td>
<td>1.8</td>
<td>4</td>
</tr>
<tr>
<td>Half-space</td>
<td>400</td>
<td>1327</td>
<td>1.8</td>
<td>Infinite</td>
</tr>
</tbody>
</table>

In the current analysis, a commonly used frequency range of 5-100 Hz (30 data points) was employed. Fig. 1 shows modeled fundamental-mode dispersion curves from all these three synthetic models.

4.1 Two-layer model

Fig. 2 demonstrates inversion results of the simple two-layer model (Model A) using the pattern search algorithm, which provide valuable insight into the performance of the proposed inversion scheme. Initial S-wave velocities are all 300 m/s for the layer and half-space, which are estimated by Eq. (1). Other inversion parameters have been mentioned in preceding section.

The convergence curve in Fig. 2a illustrates a typical characteristic of a pattern search algorithm. It shows a very fast initial convergence at the first 10 iterations, followed by progressively slower improvements as it approaches the optimal solution. Fig. 2b shows the mesh size at each iteration. The mesh size increases after each successful iteration and decreases after each unsuccessful one. The best point does not change after an unsuccessful
poll. For example, the poll at iteration 1 is unsuccessful. As a result, the algorithm halves the mesh size with contraction factor set to 0.5. We can see that the objective function value computed at iteration 2 is less than the value at iteration 1 (Fig. 2a), which indicates the poll at iteration 2 is successful. Thus, the algorithm doubles the mesh size with expansion factor set to 2 (Fig. 2b). Clearly, the poll at iteration 3 is unsuccessful. As a result, the function value remains unchanged from iteration 2 and the mesh size is halved. The inverse process is terminated after 20 iterations because the error misfit converged to approximately zero. As shown in Fig. 2c, in 20 iterations the pattern search algorithm only performs approximately 58 function evaluations to locate the promising region in the solution space containing the global minima. For this simple model, the S-wave velocities are exactly resolved (Fig. 2d). In practice, the two-layer model may be useful for estimation of static correction (Xia et al., 1999).

Fig. 2. Inversion results of Model A using pattern search algorithms. (a) The objective function value of the best point obtained at each iteration. (b) Mesh size as a function of iterations. (c) The number of function evaluations performed at each iteration. (d) Inverted S-wave velocities.

4.2 Multilayer models
The multilayer model (Model B) will help us to compare the performance of the algorithm with Model A. Initial S-wave velocities are all 250 m/s for the 4-layer model. Fig. 3 illustrates inversion results of Model B using the proposed algorithm. Similar to Fig. 2, the objective
function values improve rapidly at the first 20 iterations and then level off at the next 20 iterations (Fig. 3a). The small misfit suggests that the algorithm found the global minimum. However, it should be noted that the algorithm stops after 40 iterations. In 40 iterations the pattern search algorithm implements approximately 269 function evaluations (Fig. 3c) to search for the global optimal solution (Fig. 3d). The implementation gives the evidence that Model B is relatively more complex than Model A, as can be further verified by the fact that the maximum number (8) of function evaluations in each iterative process in Fig. 3c is greater than that (4) in Fig. 2c. Nonetheless, S-wave velocities are still accurately reconstructed (Fig. 3d).

The proposed inversion scheme is further tested on the complex multilayer model (Model C). Initial S-wave velocities are also 250 m/s for each layer of this complex model. Fig. 4 demonstrates inversion results of Model C using the pattern search algorithm. As can be seen in Fig. 4a, the misfit value decreased rapidly in the first 40 iterations and then gradually in the next 40 iterations. As expected, the algorithm stops after a larger number of iterations (80 in this inversion). In 80 iterations the algorithm carried out 677 function evaluations to discover the most promising domains for a good valley. Moreover, we should notice that the maximum number (10) of function evaluations during each implementation in Fig. 4c is greater than the number (8) in Fig. 3c, indicating that Model C is more complex than Model B. It is encouraging that S-wave velocities for all layers of Model C are still unambiguously imaged (Fig. 4d). The low velocity layer is well depicted.

Fig. 3. Inversion results of Model B using pattern search algorithms. (a), (b), (c), and (d) have the same meaning as in Fig. 2.
4.3 Effects of four different aspects on the performance of the algorithm

To evaluate the influences of four different aspects (the number of data points, a limited frequency range, errors in P-wave velocities and density, and the initial S-wave velocity profile) involved in surface wave analysis on the proposed approach, we performed four further tests on Model C using the pattern search algorithm.

For case 1, we reduce the number of data points from 30 to 15 to investigate the effect of the number of data points on the performance of the algorithm. Inversion was implemented with the same parameters previously adopted. Comparing the inversion result from 15 data points (dashed line in Fig. 5) with the true model (solid line in Fig. 5), there appears to be no decreases in accuracy because of the reduction in the number of data points and visually no changes compared with the true model. Thus, we remark that the number of data points will not have a significant influence on implementation of the proposed algorithm as long as enough data points have uniformly encountered the depths of interest. Usually, an excessive number of data points (more than needed) will require more computational time without benefit to the inversion accuracy. The result is consistent with previous findings (Rix, 1988; Xia et al., 1999).

For case 2, we performed another test by considering a limited frequency range. In practice, the frequency spectrum resulting from the application of a low-frequency source such as a sledgehammer over unconsolidated sediments often suffers from lack of high frequencies (Dal Moro et al., 2007). We have experienced this situation a number of times when
processing surface wave data. Fig. 5 shows the S-wave velocity profile inverted when only a frequency range of 5-50 Hz is used (dashed-and-dotted line). As can be seen, the solution does not significantly differ from the true model. Values in the shallower portion, however, appear somehow less precise.

Fig. 5. Effects of the number of data points, limited frequency range, errors in P-wave velocities and density, and the starting point on inversion results of Model C using pattern search algorithms.

For case 3, we introduce a 25% random error generated by a uniform distribution in both P-wave velocities and density to test the performance of the algorithm in more realistic conditions. In the real world, estimation errors in P-wave velocities and density always exist, especially for situations in which no a priori information is available. Fig. 5 illustrates the retrieved S-wave velocity profile for this case (dashed-and-dotted-dotted line). Clearly, S-wave velocities for the first three layers are well delineated. The maximum error occurs at layer 5, which is 3%. This test has also given support to our inversion choice and the suggestion by Xia et al. (1999) that it is impossible to invert Rayleigh wave dispersion curve for P-wave velocities and density.

For case 4, we study the effect of the starting point X₀ on the calculation efficiency and stability of the proposed inversion procedure. It is well known that local optimization methods are prone to being trapped at local minima, and their success depends heavily on the choice of a good starting model. Our modeling results demonstrate that pattern search methods have features that make them less likely to be trapped by spurious local minimizers than do methods that use derivatives. For this reason we are using pattern search algorithm as a global optimization technique. How does the choice of a starting point
influence the performance of the algorithm? In the following test, let us consider an extreme value 1000 m/s for all layers as an initial point at which the pattern search algorithm starts the optimization, which is approximately 300% off the maximum phase velocity of the true model. It is important to notice that the excessive choice of a starting point does not compromise the accurate reconstruction of the retrieved model and no visual differences can be found (dashed-dotted-dotted-dotted line in Fig. 5). Predictably, a much larger number of iterations (120) and more function evaluations (828) are taken for the algorithm to converge (Fig. 6). However, our testing results show that, for the extreme initial value of 1000 m/s, local optimization techniques such as Levenberg-Marquardt (Xia et al., 1999) and Occam’s algorithm (Lai et al., 2002) did not converge.

Fig. 6. Inversion results of Model C using pattern search algorithms with extreme initial S-wave velocities of 1000 m/s for all layers. (a), (b), (c), and (d) have the same meaning as in Fig. 2.

4.4 Effects of the number of layers and their thicknesses on the algorithm
In practice, the number of layers and their thicknesses may not be always a priori known in subsurface studies. In such situations it may be necessary to overparameterize the vertical profile with a larger number of thin layers. But how would the algorithm behave in this case? To further examine the performance of the algorithm in this realistic condition, we invert Model A again with a 6-layer model, each thin layer being 1.2 m thick. Initial S-wave velocities are all 300 m/s for the new 6-layer model. It is exciting that the two-layer model is still unambiguously resolved although we subdivide the two-layer subsurface into six thin layers.
Pattern Search Algorithms for Surface Wave Analysis

(Fig. 7d). As expected, the algorithm stops after a larger number of iterations (120 in this inversion) (Fig. 7a). In 120 iterations the algorithm performs 1185 function evaluations (Fig. 7c) to discover the most promising domains for a good valley. It should be pointed out that the sampling of the measured phase velocity curve provides the vertical resolution. With a large number of thin layers, a narrow sampling of the dispersion curve is necessary to resolve the S-wave velocity in each layer. We adopt a narrow sampling (120 dispersion points) to resolve the S-wave velocity in each layer in this inversion. However, we should also realize that, for any noisy data, we have to make a trade-off between the resolution and error of an inverted model, to obtain stable results. We can reduce the error in the deduced S-wave velocity model by reducing the resolution of the retrieved model (increasing thicknesses of layers).

![Graphs showing inversion results of Model A using a pattern search algorithm with a 6-layer model, each thin layer being 1.2 m thick. (a), (b), (c), and (d) have the same meaning as in Fig. 2.]

5. Effects of different inversion strategies on performance of the algorithm

GPS Positive Basis 2N and GPS Positive Basis N+1 are two commonly used inversion strategies in pattern search algorithms. They are the generalized pattern search (GPS) algorithm with the maximal positive basis set 2N vectors and the minimal positive basis set N+1 vectors, respectively, where N is the number of independent variables (which in our case involve S-wave velocities) in the optimization problem. The strategies use fixed direction vectors to define the pattern and compute the set of points forming the mesh. To investigate and evaluate effects of the above described two inversion strategies on the performance of pattern search algorithms, three synthetic earth models are used. These
models are designed to simulate situations commonly encountered in shallow engineering site investigations. As shown in Table 4, Model D represents a multilayer geologic structure with downwardly increasing S-wave velocities. Model E (Table 5) with a soft layer trapped between two stiff layers models a real complex pavement structure containing a low velocity layer. Model F (Table 6), characterized by a stiff layer sandwiched between two soft layers, simulates a multilayer ground condition containing a high velocity layer.

<table>
<thead>
<tr>
<th>Layer Number</th>
<th>S-wave velocity (m/s)</th>
<th>P-wave velocity (m/s)</th>
<th>Density (g/cm³)</th>
<th>Thickness (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>735</td>
<td>1.9</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>250</td>
<td>919</td>
<td>1.9</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>300</td>
<td>1102</td>
<td>1.9</td>
<td>6</td>
</tr>
<tr>
<td>Half-space</td>
<td>400</td>
<td>1470</td>
<td>1.9</td>
<td>Infinite</td>
</tr>
</tbody>
</table>

Table 4. Model D: A four-layer model with downwardly increasing S-wave velocities

<table>
<thead>
<tr>
<th>Layer Number</th>
<th>S-wave velocity (m/s)</th>
<th>P-wave velocity (m/s)</th>
<th>Density (g/cm³)</th>
<th>Thickness (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>611</td>
<td>1.9</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>160</td>
<td>673</td>
<td>1.9</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>260</td>
<td>1093</td>
<td>1.9</td>
<td>6</td>
</tr>
<tr>
<td>Half-space</td>
<td>400</td>
<td>1681</td>
<td>1.9</td>
<td>Infinite</td>
</tr>
</tbody>
</table>

Table 5. Model E: A four-layer model with a soft layer trapped between two stiff layers

<table>
<thead>
<tr>
<th>Layer Number</th>
<th>S-wave velocity (m/s)</th>
<th>P-wave velocity (m/s)</th>
<th>Density (g/cm³)</th>
<th>Thickness (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>120</td>
<td>441</td>
<td>1.9</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>250</td>
<td>919</td>
<td>1.9</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>1020</td>
<td>1.9</td>
<td>6</td>
</tr>
<tr>
<td>Half-space</td>
<td>400</td>
<td>1470</td>
<td>1.9</td>
<td>Infinite</td>
</tr>
</tbody>
</table>

Table 6. Model F: A four-layer model with a stiff layer sandwiched between two soft layers

Fig. 8 demonstrates inversion results of Model D using GPS Positive Basis 2N. During the inversion procedure, we set the initial mesh size at 1, expansion factor at 2, and contraction factor at 0.5. The convergence curve in Fig. 8a illustrates a typical characteristic of a pattern search algorithm. It shows a very fast initial convergence at the first 20 iterations, followed by progressively slower improvements as it approaches the optimal solution. Fig. 8b shows the mesh size at each iteration. The mesh size is doubled after each successful iteration and is halved after each unsuccessful one. The best point does not change after an unsuccessful poll. The inverse process is terminated after 36 iterations because the error misfit converged to zero. As shown in Fig. 8c, in 36 iterations the algorithm performs approximately 252 function evaluations to locate the promising region in the solution space containing the global minima. For this four-layer model, the S-wave velocities are exactly resolved (Fig. 8d). The relative errors between the inverted S-wave velocities and the true model (Table 4) are zero.
Fig. 8. Inversion results of Model D using GPS Positive Basis 2N. (a), (b), (c), and (d) have the same meaning as in Fig. 2.

Fig. 9. Inversion results of Model D using GPS Positive Basis N+1. (a), (b), (c), and (d) have the same meaning as in Fig. 2.
Fig. 9 illustrates inversion results of Model D using GPS Positive Basis N+1 with the same inversion parameters described above. As shown in Fig. 9a, the misfit value decreased rapidly in the first 80 iterations and then gradually in the next 80 iterations. The algorithm stops after a larger number of iterations (160 in this inversion). In 160 iterations the algorithm carried out 770 function evaluations (Fig. 9c) to discover the most promising domains containing the global minima (Fig. 9d).

Figs. 10 and 11 show inversion results of Model E using the above mentioned two inversion strategies with the same inversion parameters described above, respectively. Figs. 12 and 13 report inversion results of Model F using the two inversion strategies, respectively. We can draw the same conclusion as Figs. 8 and 9 by analyzing Figs. 10-13.

Comparing Fig. 8 with Fig. 9, Fig. 10 with Fig. 11, and Fig. 12 with Fig. 13, provides valuable insights into the performance of the above mentioned two inversion strategies. We understand that a pattern search might sometimes run faster using GPS Positive basis N+1 as the inversion strategy because the algorithm searches fewer points at each iteration (the maximum number of function evaluations is 5 in Figs. 9, 11 and 13c). However, as can be seen in Figs. 9, 11 and 13, the strategy, especially when it is applied to complex global optimization problems of highly nonlinear, multiparameter, and multimodal objective functions, might suffer from the high computational cost (770, 1200, and 948 function evaluations in Figs. 9, 11, and 13c, respectively) due to its slow convergence and its wandering near the global minimum in the final stage of search. In contrast, although the strategy based on GPS Positive Basis 2N can make the pattern search run longer because of exploring more points at each iteration (the maximum number of function evaluations is 8 in Figs. 8, 10 and 12c), this enables the algorithm to effectively locate the promising areas in the solution space containing the global minima by looking in more directions at each
Pattern Search Algorithms for Surface Wave Analysis

Fig. 11. Inversion results of Model E using GPS Positive Basis N+1. (a), (b), (c), and (d) have the same meaning as in Fig. 2.

Fig. 12. Inversion results of Model F using GPS Positive Basis 2N. (a), (b), (c), and (d) have the same meaning as in Fig. 2.
iteration and thereby potentially reduces the computational effort (252, 396, and 257 function evaluations in Figs. 8, 10 and 12c, respectively) due to its faster convergence at the neighborhood of the global minimum in the final stage of exploratory moves. For example, using GPS Positive Basis 2N as the inversion strategy reduces the total function count — the number of times the objective function was evaluated — by almost 68 percent. Therefore, unless otherwise noted, we will use GPS Positive Basis 2N as the inversion strategy in all of the latter tests because the strategy is preferable to that based on GPS Positive Basis N+1 for solving highly nonlinear global optimization problems.

Fig. 13. Inversion results of Model F using GPS Positive Basis N+1. (a), (b), (c), and (d) have the same meaning as in Fig. 2.

6. Effects of initial mesh size and final mesh size on performance of the algorithm

The initial mesh size $\Delta_0$ is the length of the shortest vector from the initial point $X_0$ to a mesh point. $\Delta_0$ should be a positive scalar. To scale the initial mesh size correctly, we take some experimentation using Model D. Table 7 reports effects of five different initial mesh sizes (1, 2, 4, 8, and 16) on the performance of the algorithm. For each experimentation, results from the first 8 iterations are listed. During the trials, we set the expansion factor at 2 and contraction factor at 0.5. As shown in Table 7, an excessive large $\Delta_0$ will require more computational time without benefit to the implementation of the algorithm. Clearly, the first successful poll occurs when the initial mesh size equals to 0.5. Thus, setting the initial mesh size at 0.5 should be considered good in the current study. This choice is also justified by the fact that the computation time and efficiency are improved in our latter tests.
Pattern Search Algorithms for Surface Wave Analysis

<table>
<thead>
<tr>
<th>Iter No</th>
<th>$\Delta_0 = 1$</th>
<th>$\Delta_0 = 2$</th>
<th>$\Delta_0 = 4$</th>
<th>$\Delta_0 = 8$</th>
<th>$\Delta_0 = 16$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(x)$</td>
<td>$\Delta_k$</td>
<td>$f(x)$</td>
<td>$\Delta_k$</td>
<td>$f(x)$</td>
<td>$\Delta_k$</td>
</tr>
<tr>
<td>0</td>
<td>57.74</td>
<td>1</td>
<td>57.74</td>
<td>2</td>
<td>57.74</td>
</tr>
<tr>
<td>1</td>
<td>57.74</td>
<td>0.5</td>
<td>57.74</td>
<td>1</td>
<td>57.74</td>
</tr>
<tr>
<td>2</td>
<td>36.00</td>
<td>1</td>
<td>57.74</td>
<td>0.5</td>
<td>57.74</td>
</tr>
<tr>
<td>3</td>
<td>36.00</td>
<td>0.5</td>
<td>36.00</td>
<td>1</td>
<td>57.74</td>
</tr>
<tr>
<td>4</td>
<td>36.00</td>
<td>0.25</td>
<td>36.00</td>
<td>0.5</td>
<td>36.00</td>
</tr>
<tr>
<td>5</td>
<td>10.51</td>
<td>0.5</td>
<td>36.00</td>
<td>0.25</td>
<td>36.00</td>
</tr>
<tr>
<td>6</td>
<td>10.51</td>
<td>0.25</td>
<td>10.51</td>
<td>0.5</td>
<td>36.00</td>
</tr>
<tr>
<td>7</td>
<td>10.51</td>
<td>0.125</td>
<td>10.51</td>
<td>0.25</td>
<td>10.51</td>
</tr>
<tr>
<td>8</td>
<td>10.39</td>
<td>0.25</td>
<td>10.51</td>
<td>0.125</td>
<td>10.51</td>
</tr>
</tbody>
</table>

Table 7. Effects of different initial mesh sizes on performance of pattern search algorithm.

7. Effects of expansion factor and contraction factor on performance of the algorithm

The expansion factor and contraction factor control how much the mesh size is expanded or contracted at each iteration. The multilayer model (Model D) will help us to investigate the influences of expansion factor and contraction factor on the performance of the proposed inversion procedure.

Fig. 14 demonstrates effects of expansion factor on the performance of the pattern search algorithm. During the inversion procedure, we set the initial mesh size at 0.5, which is suggested in the preceding section, expansion factor at 1, and contraction factor at 0.5, which means the pattern search multiplies the mesh size by 1 after each successful poll (i.e., not allow expansions) and multiplies the mesh size by 0.5 after each unsuccessful poll. Comparing the inversion results from Fig. 14 with the inversion results from Fig. 8, there appears to be no differences in accuracy. However, forcing the steps to be non-increasing (Fig. 14) can make the pattern search converge faster and further avoid its wandering near the global minimum in the final stage of exploratory moves. For example, the procedure with expansion factor set to 1 (not allowing expansions) reduces the total function evaluations from 252 (Fig. 8c) to 125 (Fig. 14c), which is almost 50 percent, and reduces the number of iterations from 36 to 17. Furthermore, note that at the first iteration the pattern search performs a successful poll (Fig. 14a) by scaling the initial mesh size correctly, as can be further verified in our latter tests.

Thus, we remark that the expansion steps are often a waste of computational effort. If an initial mesh size can be scaled correctly (which may take some experimentation), then there is rarely any advantage to aggressively allowing expansions. This is particularly true as the algorithm converges and the final mesh size should go to zero. As this occurs the expansion steps usually yield very little improvement and make the algorithm wander at the neighborhood of the global minimum.

Fig. 15 demonstrates how contraction factor affects the behavior of the pattern search. During the inversion procedure, we set the initial mesh size at 0.5, expansion factor at 1, and contraction factor at 0.25. As can be seen, the mesh size decreases faster with contraction factor set to 0.25 (Fig. 15b), as compared with the above value of 0.5 (Fig. 14b), and the deduced solution (Fig. 15d) does not differ from the true model. The approach, however, takes somewhat longer (157 function evaluations in Fig. 15c) to locate that best solution. Although not being addressed in this analysis, the same is true when the contraction factor is set to 0.75, as compared with the value of 0.5.
Fig. 14. Effects of expansion factor on performance of pattern search algorithm. (a), (b), (c), and (d) have the same meaning as in Fig. 2.

Fig. 15. Effects of contraction factor on performance of pattern search algorithm. (a), (b), (c), and (d) have the same meaning as in Fig. 2.
Therefore, we conclude that it is a wise strategy to continue the implementation of the algorithm with steps of the same magnitude until no further decrease in the objective function is realized, at which point the step size should be halved. This corresponds to setting expansion factor $\Lambda = 1$ and contraction factor $\theta = 1/2$. The insights issued in this section also give support to previous findings (Torczon, 1997; Lewis and Torczon, 1999).

8. Effects of inclusion of noise in surface wave data on performance of the algorithm

In practice, picked surface wave phase velocities are inevitably noisy. To further examine and evaluate effects of inclusion of noise in surface wave data on the performance of the algorithm, we introduced a 10% random error generated by a uniform distribution in surface wave phase velocities of three four-layer models (solid dots in Figs. 16, 18, and 20). Noise that had an amplitude of 10% of the phase velocity were added at each frequency. The performance of the proposed inversion scheme in this realistic condition is illustrated in Figs. 17, 19, and 21, respectively. During the inversion procedure, we adopt the parameters suggested by the above insights. We set the initial mesh size at 0.5, expansion factor at 1, and contraction factor at 0.5. Dashed lines in Figs. 16, 18, and 20 display initial S-wave velocities used in the inverse process. As shown in Figs. 17a, 19a, and 21a, the misfit values significantly decrease in the first 10 iterations, and then gradually converge to a similar constant value, which suggests that the algorithm has completed the exploration for the global minimum. As the algorithm converges and approaches the optimal solution, the final mesh size quickly goes to zero (Figs. 17b, 19b, and 21b). The inverse process is terminated after 19, 38, and 30 iterations, respectively, to prevent transferring errors in data into the desired models. Because of a correct choice for the inversion parameters, the algorithm performs only 137, 266, and 204 function evaluations (Figs. 17c, 19c, and 21c), respectively, to search for the global minima. Figs. 17d, 19d, and 21d report the best solutions of the implementation. It can be noted that

![Fig. 16. Simulation results from Model D. Solid dots, dashed line, and solid line represent contaminated, initial, and calculated fundamental-mode surface wave phase velocities, respectively.](image-url)
Fig. 17. Effects of inclusion of noise in surface wave data from Model D on performance of pattern search algorithm. (a), (b), (c), and (d) have the same meaning as in Fig. 2.

Fig. 18. Simulation results from Model E. Solid dots, dashed line, and solid line represent contaminated, initial, and calculated fundamental-mode surface wave phase velocities, respectively.
Fig. 19. Effects of inclusion of noise in surface wave data from Model E on performance of pattern search algorithm. (a), (b), (c), and (d) have the same meaning as in Fig. 2.

Fig. 20. Simulation results from Model F. Solid dots, dashed line, and solid line represent contaminated, initial, and calculated fundamental-mode surface wave phase velocities, respectively.
S-wave velocities for the first three layers of three four-layer models are well delineated, and the relative errors are all not more than 3%. Maximum errors occur at layer 4, which are approximately 4%. The modeled dispersion curves (solid line in Figs. 16, 18, and 20) from the best solutions (Figs. 17d, 19d, and 21d) also fit the measured phase velocities (solid dots in Figs. 16, 18, and 20) reasonably well.

So, we realize that pattern search algorithms possess stronger immunity with respect to noise and should be considered good not only in terms of accuracy but also in terms of computation effort when they are applied to high-frequency surface wave inversion.

![Graphs showing function value, best function value, mesh size, total function evaluations, current best S-wave velocity, and number of function evaluations over iterations.](image)

Fig. 21. Effects of inclusion of noise in surface wave data from Model F on performance of pattern search algorithm. (a), (b), (c), and (d) have the same meaning as in Fig. 2.

### 9. Comparison between pattern search algorithms and genetic algorithms

Pattern search methods proceed by conducting a series of exploratory moves about the current iteration before declaring a new iteration and updating the associated information. It is important to point out that all of the final solutions in our tests are determined by one computation instead of the average model derived from multiple trials because pattern search process is deterministic. This advantage greatly reduces the computation cost. To further highlight this feature, we use Model D again to implement a Rayleigh wave dispersion curve inversion scheme by genetic algorithms (GA) as developed by Yamanaka and Ishida (1996). According to the suggestions of Yamanaka and Ishida (1996), we set the population size at 20, crossover probability at 0.7, and mutation probability at 0.02. The algorithm is terminated at the 150th iteration. A final solution is determined from an average of 20 trials due to random constructions in a GA procedure.

Fig. 22 summarizes genetic algorithm inversion results of Rayleigh wave dispersion curves. As shown in Fig. 22a, the misfit values averaged from 20 inversions rapidly decrease in the
first 70 generations, and then gradually converge to zero in the next 80 generations, indicating that the algorithm has found the global minimum. The final solution from GA (dashed line in Fig. 22b) is identical to the true model while GA performs a total of 60,000 function evaluations in 20 trials for a final desired model. In contrast, the pattern search algorithm pursues only 252 function evaluations (Fig. 8c) to search the global minimum. Since most of the computational time is spent in the calculation of those forward problems in function evaluations, the proposed scheme applied to nonlinear inversion of surface wave

![Graph](image-url)

**Fig. 22.** Inversion results of Model D by a genetic algorithm. (a) Misfit value obtained from an average of 20 trials as a function of generation. (b) True (solid line) and inverted (dashed line) S-wave velocity profiles. The inverted model (b) is an average result from 20 inversions.
data should be considered good not only in terms of accuracy but also in terms of computation effort due to its global and deterministic search process, especially when compared to the application of genetic algorithms to surface wave inversion using the current inversion parameters.

10. Field data inversion

Modeling results presented in the previous section demonstrated the calculation efficiency and reliability of the inverse procedure. To further explore the performance of the algorithm described above, surface wave data acquired from a highway roadbed survey in Henan, China have been reanalyzed in the present study using pattern search algorithm. The surface topography of the tested roadbed is flat, without any obvious relief. The surficial layer of the roadbed has been rolled many times. A number of boreholes were used to obtain priori geologic information, and to evaluate the accuracy and efficiency of the surface wave method. Borehole SK01-06-12 (with a maximum depth of drilling of 14.5 m) reveals a 1.4-m-thick soft clay layer between a stiff surficial layer (compact sand and clay) and a gravel soil layer with mixed alluvial deposit. The soft clay layer usually leads to serious subsidence and deformation of a highway.

Based on our previous modeling result that an excessive number of data points will simply increase the computational cost without benefit to accuracy of the solution, we resample the resulting dispersion curve (Song et al., 2006) using 44 data points with a frequency range of 8-70 Hz (solid dots in Fig. 23) without compromising data accuracy. It is worth noticing that the measured dispersion curve (solid dots in Fig. 23) is characterized by an inversely dispersive trend (an abrupt variation and discontinuity) within the frequency range of 15-30 Hz, which is likely owing to the inclusion of higher modes in the fundamental mode data when a low velocity layer is present.

Similar to the inverse strategy of Model C, a 6-layer subsurface structure was adopted to perform a pattern search algorithm inversion of the observed dispersion curve. Estimated Poisson’s ratio and density are 0.4 and 1.8 g/cm³ for each layer of the 6-layer model, respectively, and are kept constant in the inverse process because the accuracy of the deduced model is insensitive to these parameters. Initial S-wave velocities approximately estimated by Eq. (1) are 250 m/s for all layers (dashed lines in Fig. 23 and Fig. 25).

The performance of the algorithm for the real example is illustrated in Fig. 24. We terminate the inversion process at a 6 m/s error level (80 iterations) to prevent transferring errors in data into the desired model. Because, in most cases, the best match with the measured data does not necessarily obtain the best inversion results (Xia et al., 2003; Song et al., 2007). Similar to Fig. 4a, the misfit values significantly decrease in the first 20 iterations, and then converge to a similar constant value (Fig. 24a), which suggests that the algorithm had completed the exploration of the good valley. In fact such behavior is common to most optimization techniques. Because of complexity of the real example, in 80 iterations the pattern search algorithm performs total 846 function evaluations (Fig. 24c) to discover the most promising areas in the solution space containing the global minima. The maximum number (12) of function evaluations during each inversion gives another evidence of complexity of the real example (Fig. 24c). Fig. 24d reports the best solution of the implementation. Clearly, the low velocity layer (the soft clay layer) has been well imaged.
The characteristic of the calculated dispersion curve from the best model is also similar to that of the modeled dispersion curve from Model C (dashed-and-dotted line in Fig. 1).

![Graph showing phase velocity vs. frequency for measured, initial, and inverted models.](image)

**Fig. 23.** A real-world example from roadbed survey in Henan, China. Solid dots, dashed line, and open circles represent measured, initial, and forwardly calculated fundamental-mode dispersion curves, respectively.

The inverted model is in acceptable agreement with borehole measurements (diamonds with a solid line in Fig. 25), especially for layer 1, layer 3, and layer 4; misfits are all not more than 4%. However, three greater misfits occur at layer 2, layer 5, and layer 6, which are approximately 8%, 5%, and 13%, respectively, although the modeled dispersion curve (open circles in Fig. 23) from the best solution (Fig. 25d) fits the measured phase velocities (solid dots in Fig. 23) reasonably well. A couple of factors, including a limited frequency spectrum, estimation errors in P-wave velocities and density, higher modes and the inclusion of noise (e.g., body waves) in the data, are likely responsible for these discrepancies. Nonetheless, the real example should be considered a successful application to nonlinear inversion of Rayleigh wave dispersion curves using pattern search algorithms. Higher modes are relatively more sensitive to the fine S-wave velocity structure than is the fundamental mode and therefore the accuracy of S-wave velocities can be further improved by incorporating higher-mode data into the inversion process (Beaty and Schmitt, 2003; Xia et al., 2003). Efforts should be made to fully exploit higher-mode Rayleigh waves for imaging and characterizing shallow subsurface more accurately, especially for the inversely dispersive site.
Fig. 24. Inversion results of the real-world example using pattern search algorithms. (a), (b), (c), and (d) have the same meaning as in Fig. 2.

11. Discussion and conclusions

From our successful inversions of synthetic and observed surface wave data, we confidently conclude that pattern search algorithms can be applied to nonlinear inversion of Rayleigh wave dispersion curves. In the current study, we are using pattern search as a global optimization technique. We remark that pattern search algorithms have a number of significant advantages compared with local optimization techniques. Firstly, the algorithms are well suited for solving problems characterized by highly nonlinear, multimodal objective functions, without the need of derivative calculations. This feature can be particularly useful to solve problems for which the objective functions are not differentiable, stochastic, or even discontinuous. However, local optimization methods may not be available in these cases. Secondly, pattern search methods have features that make them less likely to be trapped by spurious local minimizers than do methods that use derivatives. However, local optimization methods are prone to being trapped by local minima, and their success depends heavily on the choice of a good starting model. Thirdly, because pattern search algorithms require only calculating a forward model, they are easily incorporated into inversion of higher-mode Rayleigh waves as well as other geophysical inverse problems. Also, the accuracy of the partial derivatives is key in determining accuracy of the derived model. Because pattern search algorithms are based only on direct solution space
sampling, any kind of linearization of the problem to be solved is avoided, with the consequent elimination of the errors involved in such approximation. In the present work, we adopted a simple initial S-wave velocity profile and a wider searching scope, performed an extreme initial value (1000 m/s), and subdivided a two-layer subsurface (Model A) into six thin layers, to test the capability of the pattern search algorithm and to simulate more realistic cases where no a priori information is available. The results of numerical tests for both synthetic and actual field data that show that pattern search algorithms possess obvious advantages compared to local-search methods. For example, if the resulting dispersion curve suffers from a serious lack of the frequency spectrum and/or a serious contamination by the inclusion of noises (e.g., body waves) in near-surface applications, in such situations an extreme initial value and a wider searching scope may be necessary for locating a global minimum. When this occurs, local-search methods fail to correctly reconstruct the subsurface model and the initial S-wave velocities determined by Eq. (1) do not work for gradient methods, but Eq. (1) is good enough to start the pattern search algorithm.

![Graph](https://example.com/graph.png)

**Fig. 25.** Initial (dashed line), inverted (solid line), and borehole measurement (solid line with diamonds) S-wave velocity profiles.
12. Acknowledgements

This research is funded by Changjiang River Scientific Research Institute Science Foundation (Nos.YWF0906; CKSF2010009) of China. The authors extend their thanks to Professor Jianghai Xia for his constructive discussion on surface wave technique.

13. References


