

# 정합장처리에서 유전자 알고리즘에 의한 기하와 지음향 매개변수 역산

신기철\*, 박제은\*, 김재수\*\*

\* 한국해양대학교 대학원 해양개발공학과 \*\* 한국해양대학교 해양시스템공학부

## Inversion for Geometric and Geoacoustic Parameters in Matched-Field Processing

Kee-Cheol Shin\*, Jae-Eun Park\*, Jea-Soo Kim\*\*

\* Dept. of Ocean Eng., Korea Maritime Univ. Graduate School

\*\* Div. of Ocean System Eng., Korea Maritime Univ.

kcshin@kmaritime.ac.kr

본 연구는 '수중음향특화연구센터'의 지원으로 이루어졌습니다.

### 요 약

본 연구에서는 수중 음의 전달에 작용하는 기하와 지음향 매개변수들을 전역 최적화(global optimization) 방법인 유전자 알고리즘을 사용하여 추정하려한다. 유전자 알고리즘은 목적함수가 불규칙적인 경우에도 모든 가능한 매개변수들을 조사하지 않으면서 전역 최대치 또는 최소치를 추정할 수 있는 최적화 방법으로, 정합장처리와 관련된 매개변수의 조사에 적합한 방향성 조사법이라 할 수 있다. 유전자 알고리즘을 이용하여 천해 해양환경에서 수치실험을 통해 매개변수 역산의 가능성을 살펴보도록 하였다.

은 전역 최적화 방법이 가장 널리 사용된다[1-4].

음향전파모델을 사용하여 실제 음장을 모사하기 위해 요구되는 매개변수는 매우 다양하지만 계산량이나 효율성 그리고 환경에 대한 사전 정보 등으로 인해 한정적으로 사용될 수밖에 없다. 또한 가능한 매개변수들 모두 탐색하여 계산한다면 너무 많은 계산시간과 계산량이 요구된다. 따라서 이러한 문제를 해결하기 위한 방법중의 하나가 유전자 알고리즘을 이용하는 것이다. 유전자 알고리즘에 의한 매개변수 역산은 수신된 데이터와 예측된 음장을 상관시키며 다양한 매개변수들에 대해 최대 상관값을 찾는 효율적인 전역 최적화 방법이다. 본 연구에서는 유전자 알고리즘을 사용하여 해양환경에 대한 기하와 지음향 매개변수를 역산하도록 한다.

### 1. 서 론

일반적으로 수중음향 모델링 문제의 목적은 해양환경의 물리적인 매개변수를 찾는 것이다. 그러나 다중 매개변수의 역산을 수행하는 것은 근래에 이르러 컴퓨터 성능의 증가와 수치적 모델의 개발로 가능하게 되었다. 일반적으로 음장을 측정된 자료로부터 해양이나 해저층의 매개변수에 대해 직접적으로 해를 구하는 방법은 없다. 또한 매개변수의 역산은 비선형 최적화 문제가 되며, 이를 해결하는 방법은 시뮬레이티드 어닐링(Simulated Annealing, 이하 SA)과 유전자 알고리즘 같

### II. 유전자 알고리즘에 의한 역산

유전자 알고리즘은 Goldberg와 Davis에 의해 소개되었으며, Gerstoft는 수중음향 분야와 지음향 매개변수 추정 문제로 확장하여 적용하였다[3]. 유전자 알고리즘은 생물학적 진화와 유사한 기법을 기반으로 가장 효율적인 최적화 시스템중의 하나로 잘 알려져 있다. 유전자 알고리즘의 기본적인 원리는 비교적 단순하며 그림 1과 같은 과정을 거친다. 먼저, 가능한 모든 모델벡터에서 최초의 모집단을 선택한다. 각각의 표본들의 최적화는 관측 데이터와 복제 데이터 사이의 차이를 기반으로

계산한다. 진화의 단계는 개개 모델벡터의 최적화를 기반으로 초기 모집단에서 부모집단을 선택한다. 각각의 부모집단은 쌍으로 결합하여 자손을 생산하며, 이때 교배(crossover)와 돌연변이(mutation) 등으로 표현되는 유전 연산자를 적용한다. 마지막으로 우수한 형질의 자손들은 부모세대 초기의 분포와 교체되는 과정을 계속적으로 수행됨으로서 최적의 모집단을 획득한다.

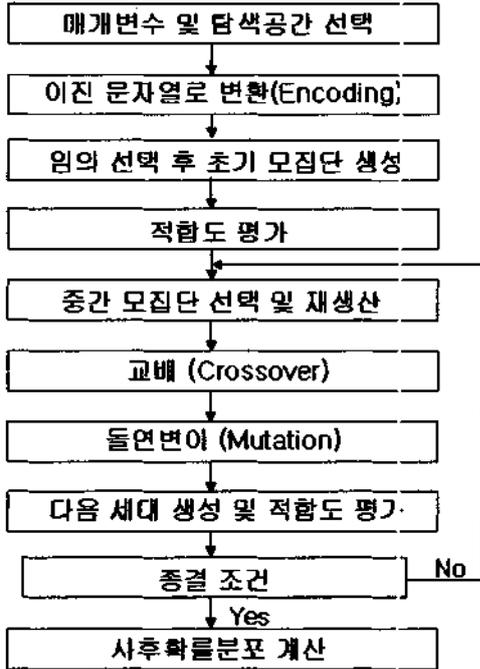


그림 1. 유전자 알고리즘의 흐름도

## 2.1 목적함수

비선형 역산 문제는 최적화 문제이며, 목적함수를 최소화하는 모델벡터  $\mathbf{m}$  를 찾는 것이다. 목적함수는 관측벡터  $\mathbf{d}$  와 음향전파모델에 의해서 예측된 벡터  $\mathbf{w}(\mathbf{m})$  의 함수이다. 목적함수는 다음의 정규화된 Bartlett 프로세서를 사용한다.

$$\phi(\mathbf{m}; w_j) = 1 - \frac{\mathbf{w}^*(\mathbf{m}) \mathbf{R}(w_j) \mathbf{w}(\mathbf{m})}{\|\mathbf{w}(\mathbf{m})\|^2} \quad (1)$$

여기에서  $\mathbf{R}(w_j)$  은 단일 주파수  $w_j$  에서 관측벡터들로부터 만들어진 공분산행렬이다. 이 목적함수를 최소화하는 것은 상관계수(correlation coefficient)를 기초로 하는 정합장 처리에서는 모호함수(ambiguity function)를 최대화하는 과정과 유사하다.

## 2.2 매개변수 초기화

해양환경은 모델벡터  $\mathbf{m}$  에 포함된  $M$  개의 환경 매개변수들로 나누어지고, 매개변수들 각각은 직사각형(rectangular) 확률분포에 따라  $2^{n_j}$  개로 이산화된다. 여기에서  $j = 1, \dots, M$  이고,  $n_j$  는  $j$  번째 매개변수 문자열(string)의 비트(bits) 수이다. 각 매개변수의 사전분포는 가우시안 확률분포와 같은 다른 형태의 분포로 선택될 수도 있다. 사전 직사각형 분포로 매개변수 각각을 이산화 했을 경우, 저계  $\theta_j^{\min}$  과 고계  $\theta_j^{\max}$  사이를  $2^{n_j}$  개의 값들로 이산화한다.

$$\theta_{j,i} = \theta_j^{\min} + \Delta\theta_j i, \quad i, j = 0, \dots, 2^{n_j} - 1 \quad (2)$$

여기에서

$$\Delta\theta_j = (\theta_j^{\max} - \theta_j^{\min}) / (2^{n_j} - 1) \quad (3)$$

프로그램에서 적합도 평가를 위해 음향전파모델 루틴이 호출되었을 때, 이산화된 모델벡터들이 사용된다. 유전자 알고리즘에서는 식(2)에서의 정수값  $i_j$  을 사용하며, 각 매개변수에 대해 모델벡터의 정수값들은 길이  $n_j$  의 이진수 문자열로 표현된다. 이러한 이진수 코딩은 교배 돌연변이 유전 연산자가 이진수 문자열에서 수행되기 때문이다.

모집단을 초기화하기 위해 탐색공간으로부터 무작위로  $q$  개의 표본 값들을 선택하고,  $k = 1, \dots, q$  의 각각에 대해 식 (1)의 목적함수를 계산한다. 각 세대에서 표본 값들은 목적함수의 값에 따라  $\phi(\mathbf{m}_1) \leq \phi(\mathbf{m}_k) \leq \phi(\mathbf{m}_q)$  와 같이 정렬된다.  $q$  개의 표본 값들을 가지고 새로운 모집단을 구성하기 위해서는  $f q$  개의 부모를 선택하여야한다. 여기에서  $0 < f < 1$  이다. 각 부모는 그들의 적합도를 기초로  $q$  개 모델들의 각각에 확률을 할당함으로써 선택된다.

SA에서 소개된 매개변수 선택을 제어하는 온도의 개념은 유전자 알고리즘에서 아주 좋은 결과를 주며, 선택의 확률은 다음과 같이 정의된다.

$$p_k = \frac{\exp[-\phi(\mathbf{m}_k)/T]}{\sum_{i=1}^q \exp[1 - \phi(\mathbf{m}_i)/T]} \quad (4)$$

여기에서  $T$  는 SA에서 정의된 것과 같은 온도이다. 지수함수의 특성으로부터 높은 온도에서는 모든 모델벡터들이 거의 동등하게 선택되는 반면에 낮은 온도에서는

목적함수의 차이가 적은 값들에 대해서도 구분이 용이하다. 그러나 낮은 온도의 단점은 가장 좋은 표본 값들만을 선택함으로써 모델의 다양성을 상실하게 된다. 이러한 두 온도 범위에서 절충안은 온도를 모집단의 목적함수와 같은 크기를 사용하는 것이다.

### 2.3 유전 연산자

부모를 선택한 후에, 새로운 세대를 구성하기 위해서 몇 가지 연산자가 응용된다. 전통적으로 이 과정은 이진수 형태로 부호화된 매개변수들에 대해 교배 연산자와 돌연변이 연산자로 구성된다.

재생산의 첫 번째 단계는 교배이다. 유전자 교배는 그림 2와 같이 한 쌍의 부모로부터 복제된 두 개의 이진 문자열(또는 유전자)을 접목하는 것이다. 모델벡터에서 하나의 매개변수에 대한 부모의 이진 문자열이  $(\alpha_0, \dots, \alpha_{N-1})$ 과  $(\beta_0, \dots, \beta_{N-1})$ 이라 할 경우에 교배점  $l$ 이  $[1, N-1]$ 로부터 무작위로 선택되면 두 자손은  $(\alpha_0, \dots, \alpha_{l-1}, \beta_l, \dots, \beta_{N-1})$ 과  $(\beta_0, \dots, \beta_{l-1}, \alpha_l, \dots, \alpha_{N-1})$ 이 된다. 이것은 각 매개변수에 대해 수행되기 때문에 이 방법을 복수점 교배라 한다.

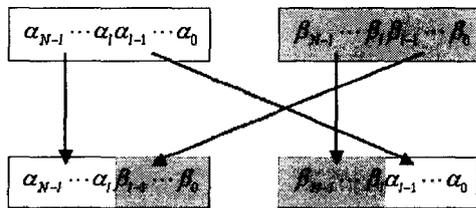


그림 2. 유전자 교배

교배를 수행한 후에, 재생산을 위한 두 번째 단계는 매개변수 벡터의 각 비트를 탐색공간에서 보다 나은 탐색을 위해 낮은 돌연변이 확률  $p_m$  ( $p_m \approx 0.05$ )로 교란시키는 것이다. 이 과정을 돌연변이 연산자라고 하며, 이는 국부적인 최소 값에 빠지지 않도록 하는데 도움을 준다. 1 비트에 대한 돌연변이 연산자의 예는 다음과 같다.

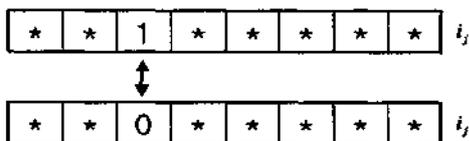


그림 3. 유전자 돌연변이

### 2.4 사후 통계

역산을 수행하는 동안 탐색공간의 표본 값들을 저장하여 사후확률(a posteriori probability)을 추정한다. Frazer 등은 각 매개변수에 대해 주변확률밀도(marginal probability density)를 표본으로 하기 위한 기법을 사용하였다. 주변확률밀도함수는 진화과정에서 모델벡터들의 표본 값에서 추정된다. 이러한 표본 값들은 목적함수의 값인 에너지에 따라 정렬되고, 확률분포를 형성할 때 Boltzmann 분포에 따라 가중되며, 이는 식 (4)와 같이 가중치가 최적화에 대해 수행되는 것과 유사하다. 표본에서 가장 적합도가 큰 에너지와 같은 온도를 선택하는 것은 모집단의 가장 적합도가 높은 부분을 선택하는 것과 같으며, 가장 적합도가 적은 에너지와 같은 온도는 모집단에 더 많은 가중치를 주는 것에 상응한다. 경험상 적절한 온도는 가장 좋은 50개의 표본을 평균하여 사용한다[3].  $k$  번째 모델벡터에 대한 확률은 다음과 같다.

$$\sigma(\mathbf{m}_k) = \frac{\exp[-\phi(\mathbf{m}_k)/T]}{\sum_{k=1}^{N_{\text{obs}}} \exp[-\phi(\mathbf{m}_k)/T]} \quad (5)$$

모델벡터에서  $i$  번째 매개변수에 대해서 특정 값  $x$ 을 얻기 위한 주변확률분포는 식 (5)을 합함으로써 얻을 수 있다.

$$\sigma(m^i=x) = \frac{\sum_{k=1}^{N_{\text{obs}}} \exp[-\phi(\mathbf{m}_k)/T] \delta(m_k^i=x)}{\sum_{k=1}^{N_{\text{obs}}} \exp[-\phi(\mathbf{m}_k)/T]} \quad (6)$$

여기에서  $N_{\text{obs}}$ 는 관측 모델벡터들의 수이고,  $T$ 는 온도이다. 또한 사후 평균값은 다음과 같이 추정된다.

$$E(\mathbf{m}) = \sum_{k=1}^{N_{\text{obs}}} \mathbf{m}_k \sigma(\mathbf{m}_k) \quad (7)$$

### III. 기하와 지음향 매개변수의 추정

매개변수 역산을 위한 기본적인 수치실험 환경은 그림 4와 같다. 즉, 주파수, 어레이 수심, 어레이 경사, 수층의 수심과 음속분포, 해저면 수심, 해저층 음속분포와 밀도 및 감쇠계수를 대상으로 수치실험하였다. 수층의 음속분포는 그림 5의 자료를 사용하였으며[5], 매개

변수 역산을 위해 사용된 상세한 매개변수의 범위는 표 1과 같다. 본 연구에서 사용된 음장의 예측 벡터는 정상모드 음향전파모델인 KRAKENC에 의해서 계산하였다[6].

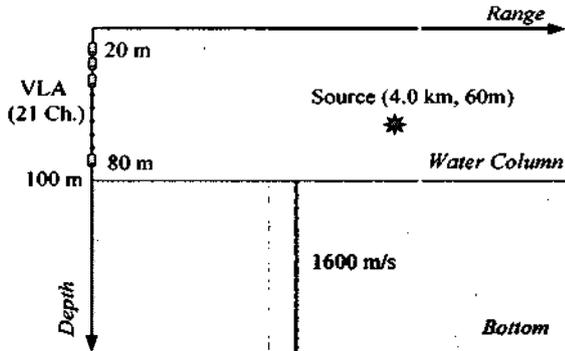


그림 4. 매개변수 역산을 위한 수치실험의 환경

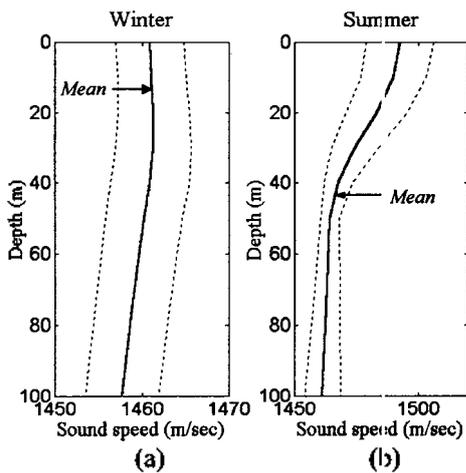


그림 5. 음속분포(실선 : 평균, 점선 : 표준편차)  
(a) 2월 (b) 8월

유전자 알고리즘을 이용한 매개변수 역산을 위해 각 매개변수 탐색공간의 크기는 128개로 하였으며, 초기 모집단의 크기는 32개로 하였으며, 재생산의 크기  $f=0.5$ , 교배율  $p_x=0.8$ , 돌연변이율  $p_m=0.05$ 로 하였다. 사후 확률분포 계산을 위한 온도는 상관계수가 높은 순서로 50개의 모집단에 대해 평균한 값을 사용하였다.

표 1. 역산을 위한 매개변수 범위

구 분	매개변수 범위		
	최소	기준	최대
음원 수심(m)	53.7	60	66.4
음원 거리(km)	3.37	4.0	4.64
주파수(Hz)	243.7	250	256.4
어레이 수심(m)	13.7	20	26.4
어레이 경사(m)	-6.3	0.0	6.4
수층 음속분포 ( $\sigma$ )	-0.63	0.0	0.64
수층 수심(m)	93.7	100	106.4
해저층 음속(m/sec)	1537	1600	1664
해저층 밀도 ( $g/cm^3$ )	1.37	2.0	2.64
해저층 감쇠계수 ( $dB/\lambda$ )	0.087	0.15	0.214

그림 6과 7은 종결 조건으로 2000회의 모델링 또는 상관 계수 0.95로 하고, 통계적인 특성을 주기 위해 50번의 독립 역산을 수행한 남해 해역에서의 2월과 8월의 수층 음속분포에 대한 매개변수별 주변확률분포이다. 그리고 표 2는 식 (7)로 계산된 매개변수별 사후 평균값이다. 2월 및 8월의 음속분포에 따른 각 해양환경에서 사후 평균값은 매개변수별로 기준되는 값과 비교해 오차가 적으나, 해저층과 관련된 매개변수와 2월의 수층 음속분포의 경우 사후 주변확률분포에서 추정값의 분산이 큼을 보였다.

#### IV. 토론 및 결론

본 연구에서는 기하와 지음향 매개변수 역산에 대해서 전역 최적화 방법인 유전자 알고리즘을 사용하여 매개변수 역산의 가능성을 조사하였다. 음향전파와 관련된 매개변수를 모두 고려했을 경우 그 조합은 계산량과 컴퓨팅 속도로 인해 불가능하다 할 수 있다. 유전자 알고리즘과 같은 전역 최적화 방법을 사용하였을 경우는 매개변수와 유사한 값을 적은 계산량으로 추정할 수 있어 역산문제에 적합한 해를 제공해 주는 것으로 보인다. 수치실험에 유전자 알고리즘을 적용하여 역산한 결과에서와 같이 실제 값과는 약간의 오차는 있었으나 허용되는 범위 내에서 신뢰성을 보였다. 특히 기하와 관련된 매개변수는 사후 확률분포에서 집중된 형태를 보였으며, 지음향 매개변수는 정합장치에서 민감하게 작용하지 않는 것으로 판단된다.

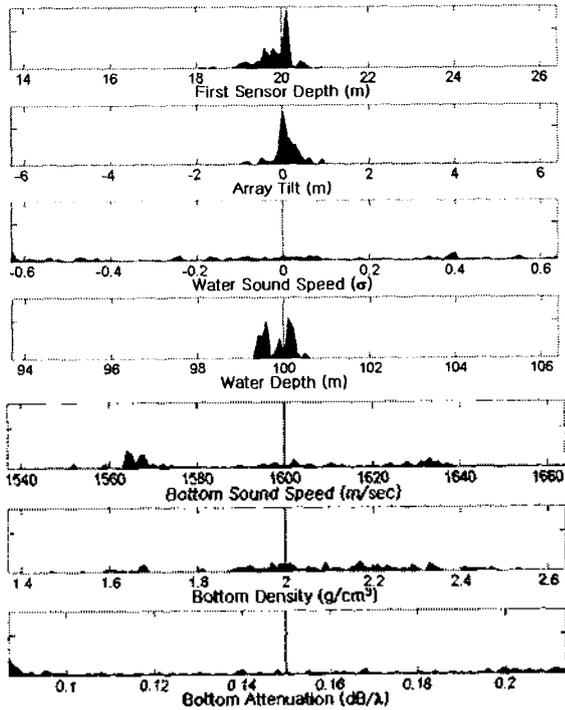


그림 6. 고정 음원에서 50회의 독립 시행 회수에 따른 2월 음속분포에서 매개변수별 주변확률분포

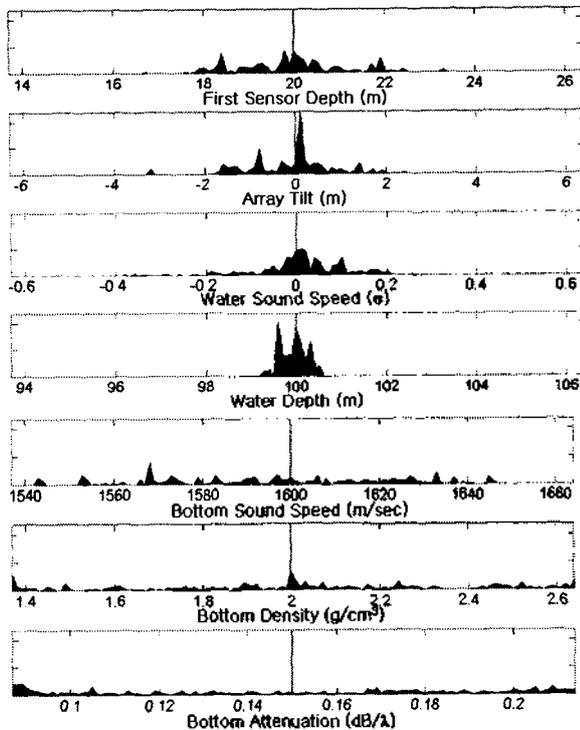


그림 7. 고정 음원에서 50회의 독립 시행 회수에 따른 8월 음속분포에서 매개변수별 주변확률분포

구분	사후 평균값		
	기준	2월	8월
어레이 수심(m)	20.0	19.8	19.8
어레이 경사(m)	0.0	0.1	0.1
수층 음속분포 ( $\sigma$ )	0.0	0.0	0.0
수층 수심(m)	100	99.9	99.9
해저층 음속(m/sec)	1600	1593.8	1599.5
해저층 밀도 ( $g/cm^3$ )	2.0	2.1	2.0
해저층 감쇠계수 ( $dB/\lambda$ )	0.15	0.1	0.1

표 2. 매개변수별 사후 평균값

### 참고문헌

- [1] M. D. Collins and W. A. Kuperman, "Focalization: Environmental focusing and source localization," J. Acoust. Soc. Am., Vol. 90, pp. 1410-1422, 1991.
- [2] M. D. Collins, W. A. Kuperman, and H. Schmidt, "Nonlinear inversion for ocean-bottom properties," J. Acoust. Soc. Am., Vol. 92, pp. 2770-1783, 1992.
- [3] P. Gerstoft, "Inversion of seismoacoustic data using genetic algorithms and a posteriori probability distributions," J. Acoust. Soc. Am., Vol. 95(2), pp. 770-782, 1994.
- [4] D. F. Gingras and P. Gerstoft, "Inversion for geometric and geoacoustic parameters in shallow water: Experimental results," J. Acoust. Soc. Am., Vol. 97(6), pp. 3589-3598, 1995.
- [5] 국립수산진흥원 한국해양자료센터 (<http://www.nfrda.re.kr/kodc/data/sodata/index-k.html>)
- [6] M. B. Porter, *The KRAKEN Normal Mode Program*, SACLANT Undersea Research Centre, 1994