

일반화된 회귀신경망과 유전자 알고리즘을 이용한 식각 마이크로 트렌치 모델링

Modeling of etch microtrenching using generalized regression neural network and genetic algorithm

이덕우, 김병환*
(Duk-woo Lee, Byungwhan Kim)

Abstract - Using a generalized regression neural network, etch microtrenching was modeled. All neurons in the pattern layer were equipped with multi-factored spreads and their complex effects on the prediction performance were optimized by means of a genetic algorithm. For comparison, GRNN model was constructed in a conventional way. Comparison result revealed that GA-GRNN model was more accurate than GRNN model by about 30%. The microtrenching data were collected during the etching of silicon oxynitride film and the etch process was characterized by a statistical experimental design.

Key Words - Plasma etching, Generalized Regression Neural Network, Genetic Algorithm, Model

1. 서 론

플라즈마 식각공정은 전자소자 제조를 위한 미세 패턴닝에 이용되고 있다. 플라즈마 특성을 예측하기 위해 신경망이 활발히 응용이 되고 있다 [1]. 일반화된 회귀 신경망 (Generalized regression neural network-GRNN) [2] 학습이 용이하고, 또한 패턴층 뉴런의 가우시안 함수가 Spread라는 단일 변수로 정의되어 있어 모델성능의 최적화가 용이하다는 이점이 있다. 종래의 GRNN의 경우 모든 학습 입력 패턴에 대해 동일한 spread를 적용하였으며, 이에 따라 패턴 간의 유사성을 구분하는 데에는 그 한계가 있었다.

본 연구에서는 종래 GRNN 모델의 단점을 극복하고자 다변수 Spread를 가지는 가우시안 함수를 정의하였으며, 다변수 spread의 모델 예측성능에의 영향을 유전자 알고리즘 (Genetic algorithm) [3]을 이용하여 최적화 한다. 본 기법을 특이 식각 특성인 마이크로트렌치 (Microtrenching) 데이터의 모델링에 적용하여 평가하였다.

2. 실 험

식각공정에 이용된 장비에 대한 구체적인 설명은 문헌 [4]에 자세히 기술되어 있다. 실험에 사용된 SiON 박막은 PECVD로 증착되었다. Ni 마스크는 박막위에 PR패턴을 형성한 후 마그네트론 스퍼터링을 이용하여 Ni층을 0.15 um

증착시킨 후 acetone으로 PR을 제거하여, Ni 마스크를 형성하였다. 그런 다음에, SiON 박막을 C₂F₆ 플라즈마로 식각하였다. 식각 공정시간은 10분으로 설정하였다.

표 1 실험변수와 범위

Parameter	Range	Unit
소스 전력	400-1000	Watts
바이어스 전력	30-90	Watts
압력	6-12	mTorr
C ₂ F ₆ 유량	30-60	sccm

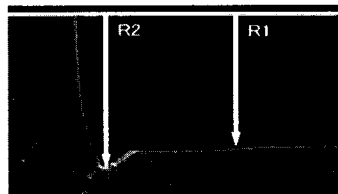


그림 1. MD를 정의하기 위한 SEM 사진

모델링을 위해 통계적인 2⁴ 인자 실험계획법 [5]가 적용이 되었으며, 관련 실험변수와 범위는 표 1에 나타나 있다. 16 번의 실험과 중앙점을 포함한 17개의 실험데이터로 모델을 학습시켰으며, 개발된 모델예측 성능을 테스트하기 위해 별도의 11번의 실험을 수행하였다. 마이크로 트렌치 깊이 (MD)는 Scanning Electron Microscope을 이용하여 수집하였다. 그림 1은 소스전력 700W, Bias 전력 60W, 압력 9mTorr, C₂F₆가스 45sccm에서 식각한 후 얻어진 SEM 영

저자 개요

* 李 德 宇 : 世宗大學 電子工學科 碩士課程

*** 金 秉 桓 : 世宗大學 電子工學科 副教授

상이며, 그림 1로부터 MD는 다음과 같이 정의된다.

$$MD = \frac{R1 - R2}{R1} \quad (1)$$

여기서 R1은 패턴간 중앙점에서의 식각률이며, R2는 프로파일 하단에서의 식각률을 의미한다.

3. 일반화된 회귀 신경망

그림 2에 도시되어 있듯이, GRNN은 총 4개의 층, 즉 입력층, 패턴층, 합층 그리고 출력층으로 구성된다. 입력층의 뉴런 수는 독립 변수의 수와 일치하며, 패턴층의 뉴런수는 학습패턴의 수와 일치한다.

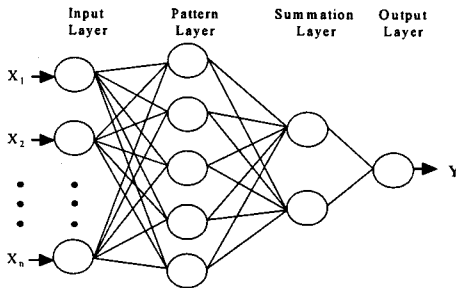


그림 2. GRNN의 구조

합 (Summation)층은 모델링 하고자 하는 각 출력변수에 대해, 2개의 S 와 D 뉴런으로 구성된다. 입력층과 패턴층간의 하중치(W_p)는 입력패턴 (X)에 의해 결정되며, 이를 표현하면,

$$W_p = X^T \quad (2)$$

여기서 "T"는 이항(transposition)을 의미한다. 패턴층의 하나의 뉴런은 합층의 두 개의 뉴런에 연결되며, 패턴층의 i번째의 뉴런과 합층 S 뉴런간의 연결 하중치는 y_i 가 된다. 여기서 y_i 는 i번째의 실험조건 (x_i)에서 측정된 증착률에 해당한다. 한편 패턴층의 i번째의 뉴런과 합층 D 뉴런간의 연결 하중치는 동일한 1값을 가진다. 따라서 연결하중치는 (3)식으로 표현된다.

$$W_s = [Y \ 1] \quad (3)$$

출력층에서의 예측치는 (4)식으로 구해진다.

$$\hat{y}_i(x) = \frac{\sum_{i=1}^n y_i \exp[-D(x, x_i)]}{\sum_{i=1}^n \exp[-D(x, x_i)]} \quad (4)$$

여기서 x_i 는 저장된 i번째의 입력 학습패턴을 지칭하며, n은 전체 학습데이터의 수를 의미한다. (4)에서 유사성을 판단하는 함수 D는 (5)식으로 정의된다.

$$D(x, x_i) = \left(\frac{x_j - x_{ij}}{\zeta} \right)^2 \quad (5)$$

여기서 P는 각 입력패턴을 구성하는 전체 독립변수의 수를 지칭한다. x_j 와 x_{ij} 는 x_j 와 x_{ij} 의 j번째의 요소를 의미한다. 그리고 변수 ζ 는 spread라 불리며, GRNN의 성능을 결정하는 유일한 학습인자이다. 일반적으로 spread는 실험적으로 일정한 범위 내에서 결정하며, 결정된 spread를 그림 2의 패턴층을 구성하는 모든 가우시안 함수가 가지게 된다.

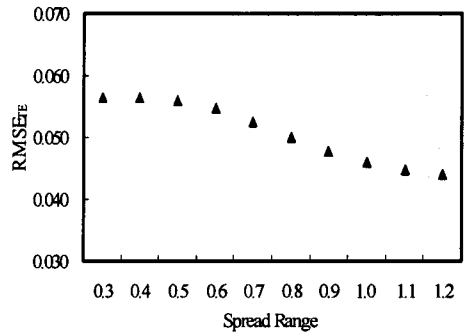


그림 3. Spread 변화에 따른 종래 GRNN의 예측성능

4. 결과

4.1 종래 GRNN 모델

비교 평가를 위해 종래의 방식으로 GRNN 모델을 개발한다. 이를 위해 설정한 spread 범위는 0.3-1.2이며, 이 범위에서 spread를 0.1간격으로 증가하였다. 각 spread에서 결정된 모델의 예측성능 (RMSE_{TE})은 그림 3에 도시되어 있다. 그림 3에서와 같이, RMSE_{TE}는 spread의 증가에 따라 감소되고 있으며, 모델 예측성능은 1.2에서 최적화되었다. 이에 해당하는 RMSE_{TE}는 0.044이다

4.2 GA-GRNN 모델

GA를 이용하여 예측모델을 개발한다. 초기 해는 300으로 설정하였으며, 각 chromosome은 총 17개의 다변수 spread

로 구성되었다. spread 범위는 0.3-1.2이며, 0.1간격으로 증가시켰다. 주어진 spread범위에서 난수발생기 (random generator)를 이용하여 17개의 spread값을 설정하였다. 교배와 돌연변이 확률은 각 0.95와 0.05으로 고정하였으며, generation 수는 100으로 설정하였다. 각 chromosome의 적합도는 (6)식을 이용하여 판정하였다.

$$F = \frac{1}{1 + RMSE_{TR}} \quad (6)$$

여기서 $RMSE_{TR}$ 는 학습 데이터로 계산한 학습에러를 의미한다.

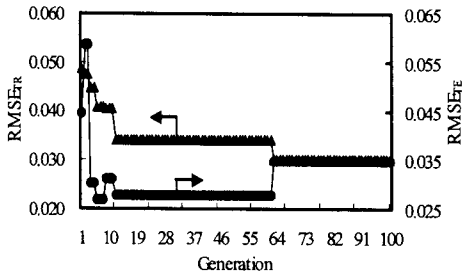


그림 4. GA-GRNN의 generation 수에 따른 $RMSE_{TR}$ 와 적합도 함수

그림 4는 spread 범위 1.0에서의 수행한 GA 최적화 과정을 $RMSE_{TE}$ 와 $RMSE_{TR}$ 로 도시하고 있다. 각 $RMSE_{TR}$ 는 매 generation 수에서 결정된 학습에러이며 $RMSE_{TE}$ 는 각 generation 수에서 결정된 최적모델 (즉 가장 작은 $RMSE_{TR}$)에 대한 예측에러이다. 그림 4에서와 같이, generation 수의 증가에 따라, 학습에러는 감소하고 있지만, 예측에러는 감소하다가 증가하는 경향을 보이고 있다. 그림 4에서, 예측에러는 6세대에서 결정된 모델이 가장 우수하며, 그 에러는 0.0268이다. 이와 동일한 방식으로 다른 spread 범위에 대해서 GA-GRNN 모델을 개발하였으며, 결정된 모델의 학습과 예측에러가 그림 5에 도시되어 있다. 그림 5에서와 같이 spread 범위가 증가함에 따라 학습에러는 증가하고 있으며, 예측에러는 증감을 보이며 감소하는 경향을 보이고 있다. 그림 5에서와 같이 가장 작은 예측에러는 spread 0.6에서 결정이 되었으며, 그 에러는 0.030이다. 종래의 모델과 비교해 볼 때, 이는 32% 정도 예측성능이 향상된 수치이다. 이는 제안하는 방식이 기존 GRNN 모델의 예측성능을 획기적으로 증진하는데 효과적임을 보여준다.

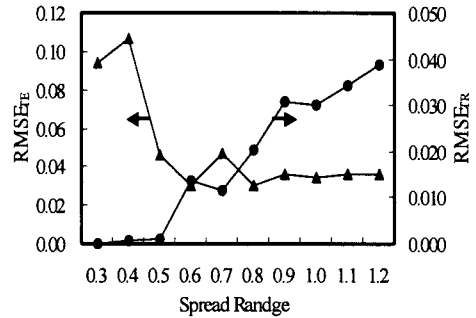


그림 5. Spread 범위에 따른 GA-GRNN 모델의 학습과 예측 에러

5. 결 론

본 연구에서 MD의 예측모델을 GRNN과 GA를 사용하여 구성하였고 그 성능을 종래의 GRNN 모델과 비교하였다. 모델개발에 이용된 공정은 SiON 박막 식각공정이며, 체계적인 데이터 수집을 위해 통계적 실험계획법이 적용되었다. 비교결과 GA-GRNN 모델은 GRNN 모델에 비해 예측성능을 향상시켰으며, 이는 다변수 spread의 최적화 방식이 단일 spread의 최적화에 비해 모델의 예측성능을 증진하는데 더 효과적임을 보여준다. 개발된 모델은 공정 최적화와 공정 변수에 따른 식각 메커니즘의 해석에 응용될 예정이다.

참고문헌

- [1] B. Kim, D. Lee, K. H. Kwon, J. Appl. Phys. 96, 3612 (2004).
- [2] D. F. Specht, IEEE Trans. Neural Network 2 (1991) 568.
- [3] D. E. Goldberg, Genetic Algorithms in Search, Optimization & Machine Learning, Addison Wesley, Reading, MA, 1989.
- [4] D. C. Montgomery, Design and Analysis of Experiments, John Wiley & Sons, Singapore, 1991.
- [5] B. Kim and B. T. Lee, J. Vac. Sci. Technol. A 22, 2517 (2004).