

# Special Thema | 플라즈마 공정의 신경망 제어기술

김병환 교수  
(세종대 컴퓨터공학부)

## 1. 서론

반도체 소자는 여러 공정을 거쳐 제조가 되며, 이들 공정 중 플라즈마를 이용한 식각과 증착공정은 박막의 증착과 미세 패턴의 형성에 활발히 이용되고 있다. 일반적으로 플라즈마 공정에는 다수의 공정 변수(예, 압력, 가스 유량)가 관여하고 있으며, 이들 변수의 미세한 변화에 따라 플라즈마 특성은 상당히 영향을 받고, 이는 다시 박막과 미세패턴에의 이상(Anomaly)를 야기해 결과적으로 소자 수율과 장비생산성을 저하시키는 주요인이 된다. 이를 방지하기 위해서는 플라즈마 공정을 엄격히 감시하고 진단하며, 동시에 제어하는 컴퓨터 기술이 요구된다. 컴퓨터 제어 시스템을 구성하는데 있어 핵심적인 역할을 수행하는 것이 플라즈마 모델이다. 플라즈마는 고전적으로 플라즈마와 박막반응현상에 대한 심도 깊은 이해와 이를 지원하는 진단 데이터에 기초하여 해석적으로 모델링이 되어 오고 있다. 해석적 모델의 경우 모델정립과정에 다수의 가정(Hypothesis)이 설정되며, 따라서 모델 예측치에의 에러가 수반될 수 있다. 모델을 구성하는 수많은 고도의 비선형적 방정식은 직접 개발한 전문가 이외에는 이해하기가 어렵고, 나아가서 공정변수나 플라즈마 소스변화에 따라 프로그램 코드를 비 전문가가 수정하는 것은 불가능하다. 한편, 모델을 계산하는데 상당한 컴퓨팅 시간이 소요되어 실시간 플라즈마 공정과 감시가 요구되는 상황에서는 적합하지 않다.

전술한 여러 문제점을 극복하기 위해 1999년 초부터 신경망이 플라즈마 공정 모델링에 응용되어 오고 있다. 신경망은 인간 뇌 작용에 대한 컴퓨터 모델이며, 통계적 회귀 모델보다 30~80% 정도의 증진된 예측성능을 보인다[1-3]. 신경망중 가장 그 응용이 활발한 구조는 역전파 신경망(Backpropagation Neural Network-BPNN)[4]이며, BPNN을 이용한 최초의 플라즈마 식각공정이 1992년 국제반도체 제조학회에서 발표된 바 있다[5]. 이후, 신경망은 다수의 플라즈마를 이용한 증착과 식각공정의 모

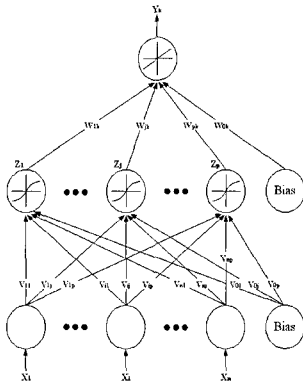


그림 1. 역전파 신경망 구조.

텔링과 해석[6-8], 최적화[9], 감시[10], 그리고 진단 [11-12]에 응용이 되어 오고 있다.

본 고에서는 먼저 플라즈마 공정제어에 활발히 응용이 되고 있는 신경망을 기본 구조와 동작원리, 그리고 모델성능을 결정하는 학습인자에 대해 설명한다. 다음에, BPNN을 이용한 플라즈마 소스 특성 모델링을 설명하고, 이어 모델링 이외의 다른 공정 감시와 진단에의 최근 연구경향을 소개한다.

## 2. 본 론

### 2.1 신경망 구조와 학습규칙

인간의 뇌는 무수히 많은 뉴런이 연결된 망 구조이며, 뉴런 상호간의 작용에 의해 정보를 처리하고 있다. 인간의 뇌 작용을 컴퓨터로 모사한 것이 신경망이며, 감독(Supervised)과 비 감독(Unsupervised) 학습 신경망으로 나눌 수 있다. 실험 데이터가 입력과 출력패턴으로 구성되어 있으면 감독학습 신경망이 이용되고 그렇지 않은 경우에는 비감독 학습 신경망이 이용된다. 플라즈마 공정 데이터는 주로 입력과 출력패턴으로 구성되며, 따라서 감독학습 신경망이 모델링에 응용되고 있다. 감독학습 신경망에는 여러 종류가 있으며, 그 중 가장 많이 응용되는 구조는 BPNN이다. BPNN의 구조가 그림1에 도시되어

있다. 그림1에서와 같이, 신경망은 3개의 층 구조로 구성되어 있으며, 각 층의 뉴런(원으로 표현)은 다른 층의 뉴런에 연결이 되고 있고, 이 들 사이의 연결강도를 웨이트(Weight)라 칭한다. 하단에 위치한 입력층에는 입력변수(소스전력, 압력, 유량 등)가 주어지며, 상단부의 출력층에는 플라즈마 또는 박막특성 변수(전자밀도, 식각률, 증착률 등)가 주어진다. 중간층은 보통 은닉층(Hidden layer)이라 칭하며, 이 층에 위치하고 있는 뉴런이 입출력간에 존재하는 임의의 비선형적인 관계를 모델링하는데 중요한 역할을 수행하게 된다. 그림 1에서와 같이 은닉층과 출력층의 뉴런은 활성화(Activation) 함수를 가지고 있으며, 대표적인 함수로는 비선형 시그모이드(Sigmoid) 함수가 있으며, 그 일반식은 식(1)로 표현된다.

$$out = \frac{1}{1 + e^{(-\frac{in}{G})}} \quad (1)$$

여기에서 in과 out은 뉴런으로의 입력과 그 뉴런으로부터의 출력을 의미하며, G는 함수의 경사를 지칭한다.

신경망 학습 과정과 규칙을 간략히 기술하고자 한다. 먼저 학습데이터를 구성하는 임의의 실험데이터의 입력 패턴이 신경망 입력단에 인가된다. 다음에 출력단에서는 인가된 입력패턴에 대한 예측치가 산출되며, 이는 실제 측정치와 비교되어 식(2)와 같이 에러가 계산되게 된다. 식(2)는 출력단 뉴런수가 하나인 경우에 계산된 것이다.

$$E = (d - out)^2 \quad (2)$$

여기에서 d와 out은 출력단 뉴런에서의 주어진 입력패턴에 대한 실제 측정치와 예측치를 의미한다. 식 (2)로 계산된 에러는 뉴런간의 웨이트의 조정에 의해 감소되며, 조정되는 웨이트 양( $\Delta W$ )는 식(3)과 같이 계산된다.

$$\Delta W = -\frac{\partial E}{\partial W} \quad (3)$$

BPNN의 학습을 지배하는 역전파 알고리즘[4]에서는 일반적으로 Gradient Descent 방식을 이용하여 식(2)의 에러가 감소된다. 은닉층과 은닉층, 그리고 은닉층과 입력층 사이에서 계산되며, 계산과정은 문헌[4]에 상세히 기술되어 있다. 계산된 을 이용하여 웨이트 정보가 식(4)와 같이 변화하게 된다.

$$W(new) = W(old) + \eta \Delta W \quad (4)$$

여기에서  $\eta$ 는 학습률을 의미하며 BPNN 학습의 수렴(Convergence)에 상당한 영향을 미치며, 대략 0.01-0.1 사이의 값을 이용한다. 식(4)로 표현되는 학습 알고리즘을 일반화된 델타 규칙[4]이라고 한다. 학습데이터를 구성하는 모든 실험의 입력패턴에 대해 이 과정이 반복된다.

### 2.2 신경망 학습인자

BPNN 학습에는 다수의 학습인자가 관여하고 있다. 학습인자에는 학습허용도, 은닉층의 뉴런수, 활성화함수의 경사, 그리고 초기웨이트 분포의 크기 등이 있다. 각 학습인자에 대한 설명과 그 영향에 대해서는 이전 연구보고에 상세히 기술되어 있다[3]. 각 학습인자는 BPNN의 학습과 예측성능에 상당한 영향을 미치며, 따라서 이들의 영향을 최적화할 필요가 있다. 학습인자 중, 그 최적화의 용이성으로 인해 은닉층의 뉴런수를 최적화에 가장 많이 이용하고 있다. 즉 일정한 실험범위 내에서 은닉층의 뉴런수를 증가시키면서 모델예측성능을 평가해 가장 작은 예측에러를 내는 모델에 대한 뉴런수를 최적의 뉴런수로 결정한다. 한편, 다른 인자들의 영향은 대부분의 모델링에서 배제되고 있으며, 우수한 예측성능을 가진 모델을 개발하기 위해서는 학습인자의 영향을 최적화하는 기법을 적용하는 것이 요구된다. 이들 학습인자 외에, 그 최적화가 가장 어려운 학습인자

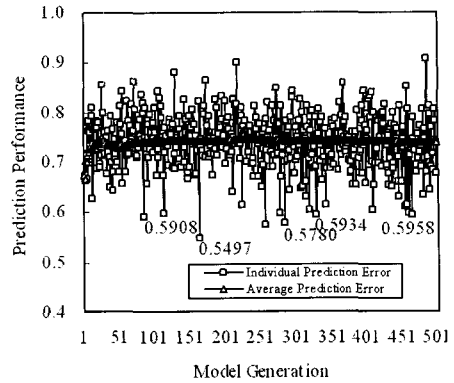


그림 2. 초기웨이트 불규칙성에 따른 모델 예측성능 분포[3].

에는 초기웨이트 분포의 불규칙성 (Randomness)이 있다. 이는 초기웨이트의 분포에 따라 모델성능이 달라지게 때문이다. 불규칙적인 초기웨이트 분포가 존재하는 환경에서, 학습인자의 영향을 해석적으로 최적화하는 것은 거의 불가능하며, 이에 따라 이를 극복하기 위한 경험적 최적화 기법이 발표된 바 있다[3]. 이 보고에서는, 통계적 실험계획법에 따라 플라즈마 방전데이터를 수집하였으며, 여기에 BPNN을 적용하여 플라즈마 모델을 개발하였다. 전술한 초기웨이트의 불규칙성을 극복하기 위해 주어진 학습인자에 대해 다수의 모델을 반복해서 개발하고 이중 예측에러가 가장 작은 모델을 선택하는 식으로 학습인자의 최적화를 이루었다. 그림2는 그 예를 도시하고 있다. 그림2는 초기웨이트의 분포에 따라 500개의 모델을 개발하였으며, 그 예측성능을 도시하고 있다[3]. 그림2에서와 같이 예측에러 분포는 초기웨이트 분포에 따라 다르며, 이는 학습인자의 영향을 최적화하는 것이 매우 어렵다는 것을 시사한다.

### 2.3 신경망 플라즈마 모델링

본 절에서는 BPNN을 이용한 플라즈마 모델링 과정을 간략히 기술한다. 소개하는 플라즈마 모델링 과정은 이전 연구보고[2]를 요약한 것이다. 신경망을 이용한 모델을 개발하기 위해서는 체계적인 방식으

로 실험데이터를 수집해야 하며, 이를 위해 통계적 실험계획법이 가장 많이 응용되고 있다. 본 예에서는 33 전인자(Full Factorial) 실험계획법을 이용하여 실험이 수행되었고, 수집된 27개의 데이터로 소위 학습데이터를 구성하였다. 테스트 데이터는 16개의 실험데이터로 구성되었으며, 주의할 점은 학습과 테스트 데이터가 서로 다른 별개의 데이터라는 점이다. 모델성능을 최적화하기 위해 이용된 학습인자에는 학습허용도, 은닉층 뉴런수, 그리고 초기웨이트 분포크기 등의 3개였으며, 이들의 영향을 개별적으로 최적화하였다. 즉 학습허용도의 영향을 최적화하고, 다음에 은닉층 뉴런수, 그리고 초기웨이트 분포 크기의 영향을 순차적으로 최적화하였다. 학습규칙으로는 전술한 일반화된 델타규칙을 적용하였다. 비교를 위해 통계적 회귀모델(Response Surface Model-RSM)을 개발하여 모델 평가에 이용하였다. 그림3은 플라즈마 전위(Potential)에 최적화된 BPNN 모델과 RSM의 테스트 실험에 대한 예측에러를 비교하고 있다.

그림3에서와 같이 BPNN 모델은 RSM 모델에 비해 대부분의 테스트 실험에 대해 실제치와 더 근접함을 보이고 있으며, 이는 BPNN 모델이 신경망 모델에 비해 예측성능이 더 우수하다는 것은 의미한다. 수치적으로, BPNN 모델은 RSM 모델에 비해 대략 35 % 정도의 증진된 예측성능을 보였다. 모델에

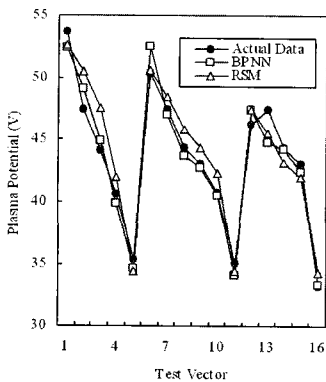


그림 3. 플라즈마 전위에 대한 모델 예측성능 비교 [2].

이용된 학습데이터 수가 작을수록 신경망과 통계모델간의 예측성능 측면에서의 격차는 더 벌어지며, 23 전인자 실험계획법으로 수집된 동일한 플라즈마 전위 모델에 대해 60% 이상의 증진된 예측성능을 보였다[3].

#### 2.4 학습인자의 최적화

BPNN 학습인자 최적화 기법에 대해 그간의 연구 동향과 결과를 소개한다. 가장 용이한 방식은, 일정한 실험범위에서 각 학습인자의 영향을 실험적으로 최적화하는 것이며, 이전 플라즈마 모델링[2]에 적용된 바 있다. 그림4는 불규칙적인 초기웨이트의 분포하에서의 초기웨이트 분포의 크기의 플라즈마 특성모델에의 영향을 도시하고 있다. 그림4에서와 같이 전자밀도, 전자온도, 그리고 플라즈마 전위 등이 모델링 되었으며, 초기웨이트 분포의 크기는  $\pm 0.2$ 에서  $\pm 1.4$ 로 변화시키면서 모델성능을 평가하고 있다. 그림4에서 최소값과 평균값은 각 초기웨이트에서 개발된 200개의 모델 중 가장 작은 예측에러를 의미하며, 평균값은 개발된 200개 모델의 예측에러의 평균치를 의미한다. 그림 4에서와 같이 각 모델에 대한 평균에러와 최소에러의 초기웨이트 분포에 따른 변화는 매우 불규칙적이다. 경우에 따라서는 평균에러 모델을 이용해 플라즈마 모델을 개발하는 경우가 있지만, 그림4에서와 같이 그 최적화된 평균에러는

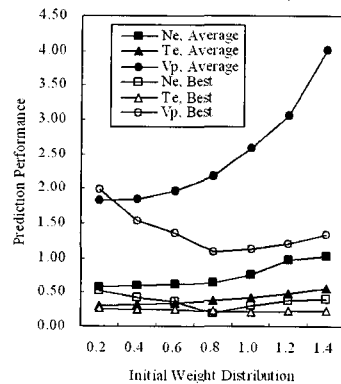


그림 4. 초기웨이트 분포의 최소에러와 평균에러 모델의 영향[3].

최적화된 최소에러보다 상당히 커, 적합한 모델링 방식이라고 할 수 없다. 평균에러 모델보다 최소에러 모델을 선택하는 것이 모델의 예측성능을 증진하는데 더 효과적이라 할 수 있다. 최소에러 모델에 기초한 학습인자의 최적화에 대해서는 기존 연구보고 [3]에 상세히 기술되어 있다. 이 방식에서는 학습인자의 영향을 최적화하기 위해 통계적 실험계획법이 적용되었으며, 여기에 유전자 알고리즘(Genetic Algorithm-GA)이 적용되어 최적화된 학습인자를 구하였다. 하지만, 이 방식에서는 학습인자 영향 모델개발에 수반되는 에러와 GA 최적화에 수반되는 에러로 인해 최적화 결과의 신뢰도가 저하될 수 있다는 위험이 있다. 이를 극복하기 위해서는 학습인자 영향모델을 개발하지 않고 GA를 직접 학습인자에 적용한 최적화 기법이 요구되며, 최근 관련 연구결과[13]가 보고된 바 있다. 이 방식에서는 학습인자 영향모델의 개발이 없이 다수의 학습인자에 GA를 직접 적용하여 모델예측성능을 최적화하였으며, 고전적인 최적화 방식(즉 개별 학습인자의 순차적 최적화)과 비교하여 평균 20% 이상의 증진된 예측성능을 보였다. 한편, 뉴런 활성화 함수의 경사를 실험적으로 최적화할 경우, 모든 뉴런의 활성화 함수가 동일한 경사 값을 가지게 된다. 만약 은닉층과 출력층 뉴런의 활성화 함수의 경사 값이 서로 다른 값을

가질 때, 모델성능의 상당한 변화를 예측할 수 있으며, 이는 그림5에 잘 나타나 있다[14]. 그림5에서와 같이 뉴런 활성화 함수 경사 조합에 따라 모델의 예측성능이 상당히 변화하고 있다. 그러나, 이 방식에서는 비록 은닉층과 출력층 뉴런의 경사값이 다르더라도 각 층의 경사값은 동일하다는 한계를 지니고 있다. 이를 해결하기 위해 각 층의 경사 값을 서로 다른 변수로 설정하고 GA를 적용하여 최적화된 경사값을 구하였으며, 그 결과 예측성능을 33-83% 정도 증진시켰으며[14], 이는 다변수 경사가 우수한 예측성능을 가진 모델을 개발하는 데에 효과적임을 보여 준다.

이외에, 아직 발표되지는 않았지만, 본 연구실에서는 다변수 경사 모델과 이를 제외한 다른 학습인자를 동시에 최적화하는 기법을 개발하였으며, 단순한 다변수 경사모델에 비해 증진된 예측성능을 확인하였다. 이상에서와 같이 최적화 기법을 적용한 모델은 적용하지 않은 모델에 비해 20-80% 정도 예측성능이 증진될 수 있으며, 따라서 BPNN을 이용한 모델개발을 하는 데 있어 학습인자를 최적화하는 기법이 적용되는 것이 요구된다.

### 2.5 신경망 응용 플라즈마 감시와 진단

신경망은 플라즈마 감시와 진단을 위해 응용이 되고 있다. 신경망을 시계열 모델링 기법에 적용하여 공정변수나 챔버내 실시간 플라즈마 진단 변수를

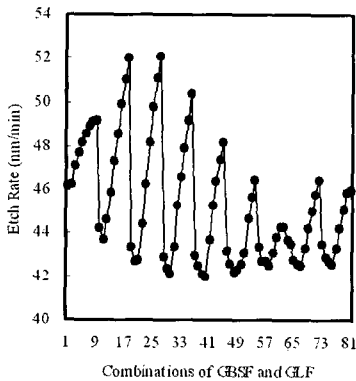


그림 5. 뉴런 활성화함수 경사값의 모델 예측성능에의 영향 [14]

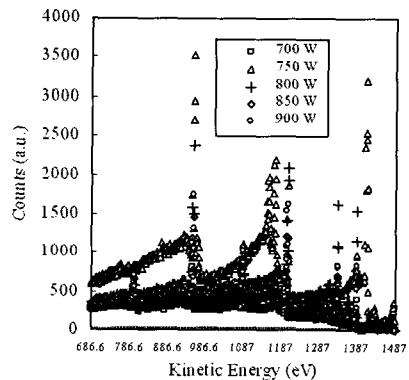


그림 6. 소스전력변화에 따른 XPS 패턴 변화.

예측하고 이를 기초하여 플라즈마 고장의 조기 감시와 진단을 수행되고 있다. 한편, 신경망은 공정 후 수집되는 박막표면상의 이상정보에 기초해서 고장 원인을 진단하는 진단기 설계에도 응용된 바 있었지만, 제한된 수의 표면정보로 인해 진단성능을 증진하는데에는 한계가 있었다[11]. 신경망 진단기 설계를 위해서는 고장 입력 패턴이 가능한 많은 변수로 표현되어야 하며, 이를 위해서는 Scanning Electron Microscope (SEM), Atomic Force Microscopy (AFM), X-ray Photoelectron Spectroscopy(XPS) 등이 진단에 유용한 미세 정보를 제공해 줄 수 있다. XPS 정보는 표면상의 화학적 상태에 대한 미세정보를 제공하며, 신경망과 결합한 진단기술이 본 연구실에서 최근 개발된 바 있다. 그림6은 Radio frequency 소스전력을 변화시켜가면서 수집한 XPS 패턴의 변화를 보이고 있다. 그림6에서 알 수 있듯이 XPS 패턴은 소스전력에 따라 특이한 패턴을 보이고 있으며, 이는 플라즈마 이상을 탐지하고 진단하는데 활용할 수 있다. 본 연구실에서는 최근 XPS 패턴의 신경망 인식 기술에 대해 연구하였으며, 수집된 XPS 패턴에 대한 고장원인을 정확하게 진단하는 것을 확인할 수 있었다. 최근 SEM 정보는 웨이브릿 기법과 연계하여 플라즈마 감시에 응용된 바 있으며[15], 현재 신경망과 결합한 플라즈마 진단기술이 본 연구실에서 연구 중에 있다. 한편 AFM에 의해 측정되는 박막표면상의 표면거칠기 패턴은 플라즈마 이상을 탐지하는데 활용할 수 있으며, 현재 웨이브릿과 신경망을 결합하여 관련 감시와 진단기술을 본 연구실에서 개발하고 있다.

### 3. 결론

본 고에서는 신경망의 플라즈마 공정모델링과 감시, 그리고 진단에의 응용에 관하여 살펴보았다. 신경망은 장비센서정보와 박막표면정보의 인식과 예측에 탁월한 성능을 보여오고 있다. 이러한 신경망 성능을 반도체 플라즈마 공정 중에 수집되는 다양한

정보와 결합하여 감시와 진단용 컴퓨팅기술을 개발할 때 장비생산성의 증대와 소자신뢰성의 확보가 기대된다.

### 참고 문헌

- [1] C. D. Himmel and G. S. May, "Advantages of plasma etch modeling using neural network over statistical techniques," IEEE Trans. Semicond. Manufact., Vol. 6, No. 2, p. 103, 1993.
- [2] B. Kim and G. T. Park, "Modeling plasma equipment using neural networks," IEEE Trans Plasma Science, Vol. 29, No. 1, p. 8, 2001.
- [3] B. Kim and S. Park, "An optimal neural network plasma model: a case study," Chemometr. Intell. Lab. Syst., Vol. 56, p. 39, 2001.
- [4] D. E. Rummelhart and J. L. McClelland, Parallel Distributed Processing, Cambridge, M.I.T. Press, 1986.
- [5] C. Himmel, B. Kim, and G. May, "A Comparison of Statistically-based and Neural Network Models of Plasma Etch Behavior," Proc. 4th Internat'l Semi. Manufa. Sci. Symp., p. 124, 1992.
- [6] B. Kim, K. Park, and D. Lee, "Use of neural network to model a deposition rate of PECVD-silicon nitride films," Plasma Source Sci. Technol., Vol. 14, p. 83, 2005.
- [7] B. Kim, D. Lee, K. H. Kwon, "Prediction of etch microtrenching using a neural network," J. Appl. Phys., Vol. 96, No. 7, p. 3612, 2004..
- [8] B. Kim and K. Kwon, "Qualitative modeling of silica plasma etching using neural network," J. Appl. Phys., Vol. 93, No. 1, p. 76, 2003.
- [9] S. S. Han, L. Cai, G. S. May, A. Rohatgi, "Modeling the Growth of PECVD Silicon Nitride Films for Solar Cell Applications Using Neural Networks," IEEE Trans. Semicond. Manufact., Vol. 9, No. 3, p. 303, 1996.
- [10] M. D. Baker, C. D. Himmel, and G. S. May, "Time series modeling of reactive ion etching using neural networks," IEEE Trans. Semicond. Manufact., Vol. 8, No. 1, p. 62, 1995.
- [11] B. Kim and G. S. May, "Real-time diagnosis of semiconductor manufacturing equipment using a hybrid neural network expert system," IEEE

Trans. Comp. Packag. Manufact. Technol., Vol. 20, No. 1, p. 39, 1997.

[12] S. J. Hong and G. S. May, "Neural network-based real-time malfunction diagnosis of reactive ion etching using in-situ metrology data, IEEE Trans. Semicond. Manufact., Vol. 17, No. 3, p. 408, 2004.

[13] B. Kim and J. Bae, "Prediction of plasma processes using neural network and genetic algorithm," Solid-State Electronics, Vol. 49, p. 1576, 2005.

[14] B. Kim, S. Kim, "GA-optimized backpropagation neural network with multi-parameterized gradients and applications to predicting plasma etch data," Chemometr. Intell. Lab. Syst., Vol. 79, p. 123, 2005.

[15] B. Kim, W. Choi, and M. T. Lim, "Wavelet monitoring of plasma etching," J. Vac. Sci. Technol B, Vo. 21, No. 6, p. 2329, 2003.

저|자|약|력



성 명 : 김병환

◆ 학 력

- 1985년 고려대 전기공학과 공학사
- 1987년 고려대 전기공학과 공학석사
- 1995년 Georgia Institute of Technology 전기공학과 공학박사

◆ 경 력

- 1996년 - 1998년 (주)현대전자 책임연구원
- 1999년 - 2001년 전남대 전임강사
- 2001년 - 현 재 세종대 컴퓨터공학부 부교수

