

통계기법을 이용한 회분식공정에서의 이상진단

정기택*, 이창준, 송상옥 윤인섭
 서울대학교 응용화학부
 (hans1611@hotmail.com*)

A Study on Process Fault Detection Systems Based on Multy-way Methods

Ki Taeg Jung*, Chang Jun Lee, Sang Ok Song, In Sup Yoon
 School of Chemical Engineering, Seoul National University
 (hans1611@hotmail.com*)

배경

많은 화학공정과 같이 반도체 공정에서도 이용할 수 있는 다양한 센서들과 센서 시스템들로 인해서 모든 시간에 있어서 많은 측정 데이터를 가지게 되고 있다. 추가된 센서들의 목적은 공정의 제어 및 이상 진단을 개선함으로써 최종 생산물의 질을 향상함으로써 비용을 줄이고자 함이다. 그러나 센서들은 종종 그것들의 목적과는 다른 경우가 많다. 센서가 유용하기 위해서는 각각의 센서들은 공정의 변화에 민감해야 하며, 오랜 시간에 걸쳐 안정된 정보를 충분히 제공되어야 한다. 뿐만 아니라, 공정 데이터를 다루는 방법들은 민감성과 강건성을 수행할 수 있도록 특성화되어져 있어야 한다.

이상 진단은 온라인으로 공정정보를 분석, 관리하고 공정의 상태를 감시하여 이상이 발생하였을 때 이상의 원인을 진단하고, 이상의 계속적인 전파를 방지하여 손실을 줄일 수 있도록 조업자의 판단과 의사결정에 필요한 정보를 제공하여 공정 이상에 효율적으로 대처할 수 있도록 도와주는 역할을 수행한다. 여기서 이상이란 외부로부터 전해지는 비정상적인 외란 (external disturbance), 조업에서 발생할 수 있는 조업자의 과실(human error), 장치의 물리적인 고장 (hardware failure) 등 공정을 정상적인 조업상태에서 벗어나게 만드는 원인을 말한다.

반도체 공정의 batch 공정에 대한 이상감지 및 진단을 위해 multi-way 통계기법들을 이용하는 많은 연구들이 행하여져 왔고, 이 방법들에 대한 비교도 이루어지고 있다. 이 기법들은 각 공정 데이터의 특성을 고려하여져 선택 되어져야 하며, 반도체 공정 중 etch 공정에서의 데이터를 이용하여 이 방법론들의 이상 감지 성능을 비교하였다. 여기서 문제는 batch 공정의 특성상 공정 및 센서의 이동을 어떻게 다루느냐 하는 것이다. 이는 센서가 이상을 감지하는데 있어서 가장 큰 영향을 줄 수 있기 때문에 이에 대한 문제가 해결 되어야만 보다 정확한 이상 진단이 이루어 질 수가 있다.

공정의 이상 실험

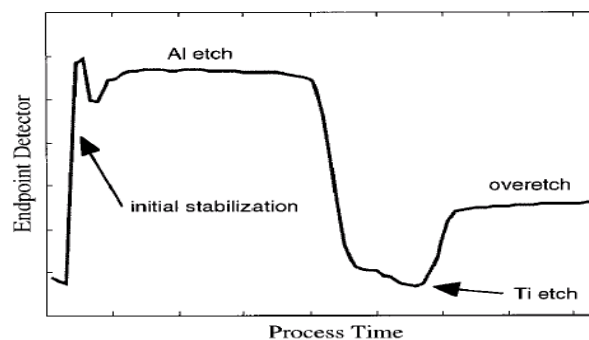


Figure 1. Endpoint trace

대상 공정은 Lam 9600 플라즈마로 Al을 식각하는 공정이며, 이 공정의 목적은 BC13/Cl2 플라즈마를 이용하여 TiN/Al-05% Cu/TiN/oxide 로 식각하는 것이며, 여기서의 주요 파라미터는 식각된 Al의 라인의 길이 폭이다. 이 공정은 여섯 단계로 이루어져 있으며, 이는 그림 1에서 endpoint의 프로파일을 보여주고 있다.

이 공정에서 이상 진단의 실험을 하기 위해서 임의적으로 fault를 넣어서 이상 진단 성능을 확인하였다. 이에 사용된 fault는 TCP power, RF power, pressure, Cl2 and BC13 flow rate 등이며, 총 129개의 웨이퍼 중 21에 웨이퍼에 이상에 가해졌다. 측정 변수는 모두 21이다.

다변량 통계 기법

1) MPCA(multi-way Principal Component Analysis)

그와 같은 원리 인 MPCA는 배치공정 데이터를 모니터링 및 분석에 사용된다. 배치공정은 배치 수, 측정변수 및 시간으로 구성되어 있다. 데이터는 그림 1. 와 같이 I JK로 구성된 삼차원 데이터 행렬로 구성된다. 삼차원 배열(I JK)를 X 로 두자. I는 각기 다른 배치, J는 측정된 변수를 그리고 K는 측정 시간을 나타낸다고 한다. MPCA는 3차원으로 구성되어 있는 데이터에 사용되는 방법론 중에 하나이다.

$$X = \sum_{r=1}^R t_r \otimes P_r + E$$

여기에서 \otimes 는Kronecker product를 나타내며, R는 주성분의 수, t_r 는 배치 사이의 관계, P 은 시간의 변화에 따른 측정변수 값 그리고 E는 오차 행렬을 나타낸다. 여기서 E는 최소자승법에 의해 최소값을 가지도록 설정 한다.

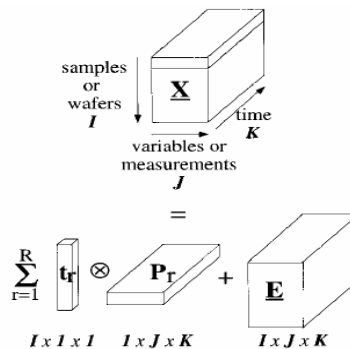


Figure2. Three-dimensional data array and multi-way PCA decomposition

MPCA 방법은 가능한 여섯 가지 방법 중(이 중 수학적으로 가능한 것은 세가지 뿐이다.)에 하나로 그림 2.와 같이 3차원 배열의 X 를 펼쳐서 큰 2차원 배열 행렬로 만든 다음 PCA를 수행하는 것과 같다.

- Batches variables at each time (time-wise unfolding)
- Variables times for each batch (batch-wise unfolding)
- Batches times for each variable (variable-wise unfolding)

Time-wise 2차원 배열은 샘플들 사이에서의 변수를 분석하는데 사용되며, batch-wise는 시간의 변화에 따라 측정된 변수들과 관계된 정보를 요약해서 이를 이용해 배치 변화를 분석하는데 이용된다. 마지막으로 variable-wise는 각 배치 변수들 사이의 변화에

대한 정보를 얻고자 할 때 이용된다.

2) PARAFAC(Parallel Factory Analysis)

MPCA는 2차원의 PCA에 의존하며, 단순히 원본 데이터의 재정렬을 하는 기술로 'poor man's' multi-way 기술이다. 이에 추가적으로 loading matrices는 변수의 정보 및 시간이 복잡하게 얽혀져서 해석하기가 어려운 문제가 있다. 이에 PARAFAC는 원본 배열을 기존의 차원 각각으로 분해하는 multi-way 방법이다. 이 모델은 아래와 같은 식으로 나타낸다.

$$d_{ijk} = \sum_{r=1}^R a_{ir} b_{jr} c_{kr} + e_{ijk}$$

여기서 R는 e_{ijk} 의 요소를 가지는 행렬 E가 작은 값을 가지게 선택되는 값이다.

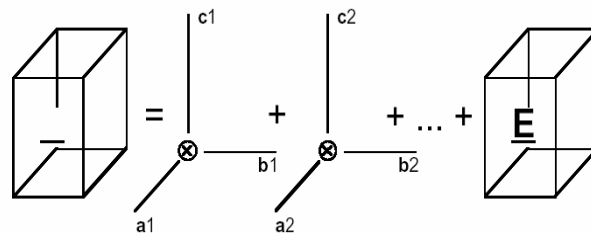


Figure 3. PARAFAC Model

여기서 A, B, C 각각의 행렬에서 r번째 열이 a_r, b_r, c_r 로 정의된다면, a_r, b_r, c_r 의 외산(outer product)은 PARAFAC의 r번째 모델이 된다. 즉, PARAFAC 모델은 세개가 한별이 되는 벡터의 출력물의 전체 합으로 나타나게 된다. 이는 그림 3.에서 나와 있는 것과 같다. 이는 PCA와는 달리 이 모델은 모든 차원에서의 loading을 제시하는 것이 특징이다.

3) FDA(Fisher Discriminant Analysis)

FDA는 선형의 차원을 줄여주는 방법으로, 이는 각각의 클래스간의 흠어짐을 최대로 함과 동시에 각각 클래스 내의 흠어짐을 최소화하는 선형의 변환 벡터의 집합으로 최적화하는 기술이다.

여기서 n은 관측한 수, m은 측정변수의 수이며, p는 클래스의 수로 정의하면, n_j 는 j번째 클래스에서 관측한 수를 가리킨다. 또한 i번째 측정 변수의 벡터를 x_i 로 나타낸다.

$$S_t = \sum_{i=1}^n \left(x_i - \bar{x} \right) \left(x_i - \bar{x} \right)^T$$

이는 total-scatter matrix 이다. 그리고 여기서 \bar{x} 는 전체 평균 벡터를 나타낸다.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

x_j 는 j번째 클래스에 속하는 x_i 벡터의 집합으로 정의되며, j번째 클래스에 있어서 within-scatter matrix는 아래와 같다.

$$S_j = \sum_{x_i \in X_j} \left(x_i - \bar{x}_j \right) \left(x_i - \bar{x}_j \right)^T$$

\bar{x}_j 는 j번째 클래스의 평균 벡터를 나타낸다.

$$\bar{x}_j = \frac{1}{n_j} \sum_{x_i \in X_j} x_i$$

within-class-scatter matrix는 아래와 같은 식으로 나타난다.

$$S_w = \sum_{j=1}^p S_j$$

그리고 between-class-scatter matrix는 아래와 같다.

$$S_b = \sum_{j=1}^p n_j \left(\bar{x}_j - \bar{x} \right) \left(\bar{x}_j - \bar{x} \right)^T$$

total-scatter matrix는 between-class-scatter matrix와 within-class-scatter matrix의 합과 같다.

$$S_t = S_b + S_w$$

FDA 첫번째 벡터의 목적은 클래스간의 거리를 최대화 함과 동시에 각각 클래스 내의 거리를 최소화는 하는데 있다.

$$\max_{v \neq 0} \frac{v^T S_b v}{v^T S_w v}$$

FDA 두번째 벡터는 클래스 간의 거리를 최대화함과 동시에 첫번째 FDA 벡터에 수직 축으로 투영하였을 때 각각 클래스 내의 거리가 최소가 되게 계산이 된다.

결론

반도체 etch 회분식 공정에서의 multi-way 통계기법을 사용하여 이상 진단을 수행하였다. 이에 사용한 다양한 통계 기법들을 비교함으로써 각 공정의 특성에 따라 기법들의 성능을 비교 하였다. 여러 논문들을 보면 알 수 있듯이 아직까지 어느 multi-way 기법의 성능이 가장 우수하다고 할 수는 없으나, 이에 사용된 공정에서는 PARAFAC이 가장 좋은 성능을 보였으며, FDA의 방법도 회분식 공정에서의 진단 성능을 확인 할 수가 있었다.

회분식 공정에서의 가장 큰 문제는 서론에서 제안한 것과 같이 batch 공정의 특성상 공정 및 센서의 이동을 어떻게 다루느냐 하는 것이다. 이는 센서가 이상을 감지하는데 있어서 가장 큰 영향을 줄 수 있기 때문에 이에 대한 문제가 해결되어야만 보다 정확한 이상 진단이 이루어 질 수가 있다.

감사

본 연구는 산업자원부/정보통신부 지원의 “화학 산업의 웹기반 지능형 온라인 통합생산 관리시스템”과제 및 교육인적자원부의 Brain Korea 21의 지원에 의한 것입니다.

참고문헌

1. L. H. Chiang, E. L. Russell and R. D. Braatz, *Fault Detection and Diagnosis in Industrial System*, Springer, 2001.
2. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, 1973
3. Barry M. Wise, *A comparison of principal component analysis, multyway principal component analysis, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process*, Journal of Chemometrics, 1999