



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2021년07월16일  
(11) 등록번호 10-2279045  
(24) 등록일자 2021년07월13일

(51) 국제특허분류(Int. Cl.)  
G05B 23/02 (2006.01) G05B 13/02 (2006.01)  
H01L 21/67 (2006.01)  
(52) CPC특허분류  
G05B 23/0294 (2013.01)  
G05B 13/0265 (2013.01)  
(21) 출원번호 10-2019-0096657  
(22) 출원일자 2019년08월08일  
심사청구일자 2019년08월08일  
(65) 공개번호 10-2021-0017431  
(43) 공개일자 2021년02월17일  
(56) 선행기술조사문헌  
JP2003158108 A\*  
JP2010027701 A\*  
KR100572039 B1\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
연세대학교 산학협력단  
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)  
(72) 발명자  
김창욱  
서울특별시 서대문구 연세로 50 연세대학교 제4공학관 1016호  
김신영  
서울특별시 서대문구 연세로 50 연세대학교 제4공학관 1012호  
(74) 대리인  
특허법인우인

전체 청구항 수 : 총 16 항

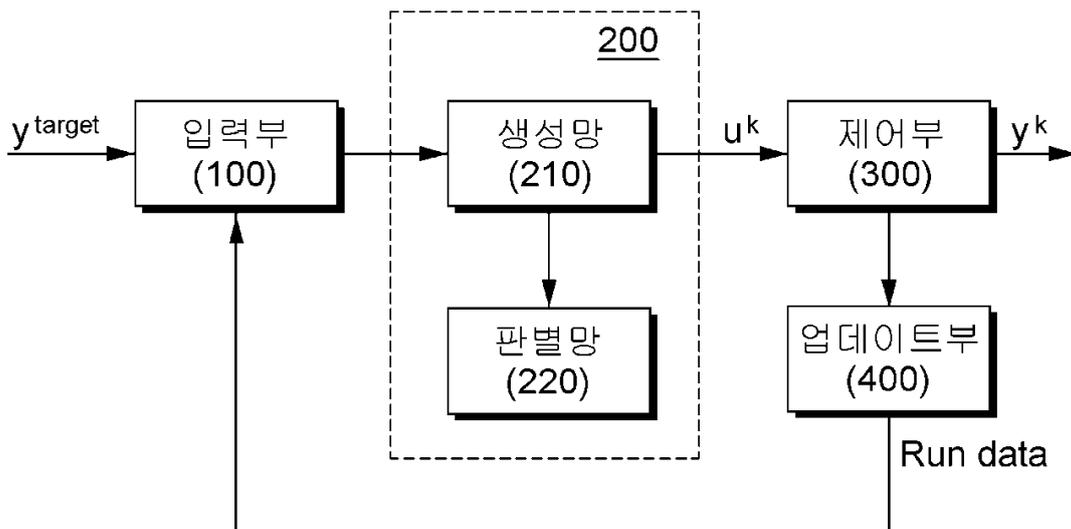
심사관 : 김윤한

(54) 발명의 명칭 공정 제어정보 생성 장치, 방법 및 이를 포함하는 공정 제어장치

(57) 요약

본 발명의 실시예에 따른 기관 제조공정에 필요한 공정 제어정보 생성 장치는, 제조하고자 하는 기관에 속하는 영역들의 스케일과 관련된 목표 스케일 정보 및 상기 기관을 제조하기 위한 적어도 하나의 장비에 대한 공정 파라미터 정보를 입력 받는 입력부, 상기 목표 스케일 정보 및 상기 공정 파라미터 정보를 기초로 복수개의 레이어들을 이용하여 상기 기관을 제조하기 위한 공정에 필요한 제1 공정 제어정보를 생성하는 생성망 및 상기 제1 공정 제어정보 및 트레이닝 데이터인 제2 공정 제어정보를 기초로 학습함에 따라 상기 제1 공정 제어정보에 대한 평가 결과를 도출하는 판별망을 포함할 수 있다.

대표도 - 도2



(52) CPC특허분류

G05B 23/0224 (2013.01)  
 H01L 21/67253 (2013.01)  
 H01L 21/67276 (2013.01)  
 G05B 2219/45232 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	NRF-2019R1A2B5B01070358
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구자지원사업
연구과제명	인공지능 스케줄러 기반 지능형 생산 관리 시스템
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2019.06.01 ~ 2022.02.28
공지예외적용	: 있음

---

**명세서**

**청구범위**

**청구항 1**

제조하고자 하는 기관에 속하는 영역들의 스케일과 관련된 목표 스케일 정보 및 상기 기관을 제조하기 위한 적어도 하나의 장비에 대한 공정 파라미터 정보를 입력 받는 입력부;

상기 목표 스케일 정보 및 상기 공정 파라미터 정보를 기초로 복수개의 레이어들을 이용하여 상기 기관을 제조하기 위한 공정에 필요한 제1 공정 제어정보를 생성하는 생성망; 및

상기 제1 공정 제어정보 및 트레이닝 데이터인 제2 공정 제어정보를 기초로 학습함에 따라 상기 제1 공정 제어정보에 대한 평가 결과를 도출하는 판별망;을 포함하며,

상기 생성망은,

상기 기관의 공정 제어와 관련되어 기 저장된 공정 제어 히스토리 정보를 이용하여 제1 특징 맵을 생성하고, 상기 제1 특징 맵을 서로 다른 스케일의 필터들을 각각 이용하여 컨볼루션 필터링 함에 따라 드리프트(drift) 특징정보를 생성하는 드리프트 특징정보 생성부;를 포함하는 기관 제조공정에 필요한 공정 제어정보 생성 장치.

**청구항 2**

제1항에 있어서,

상기 스케일 정보는 상기 기관에 속하는 각 영역의 두께 정보이고,

상기 공정 파라미터 정보는 상기 장비의 누적 사용시간에 대한 정보이며,

상기 제1 공정 제어정보 및 제2 공정 제어정보는 단일 운용과정에서 소요되는 폴리싱 시간 및 폴리싱 압력 값을 포함하는 것을 특징으로 하는 기관 제조공정에 필요한 공정 제어정보 생성 장치.

**청구항 3**

삭제

**청구항 4**

제1항에 있어서, 상기 생성망은,

상기 드리프트 특징정보, 상기 스케일 정보들 및 상기 공정 파라미터 정보를 입력 받아 적어도 하나의 컨볼루션 레이어를 이용하여 컨볼루션 필터링하여 FC 특징정보를 생성하는 FC 특징정보 생성부;를 더 포함하고,

상기 FC 특징정보를 디코딩하여 상기 제1 공정 제어정보를 출력하는 것을 특징으로 하는 기관 제조공정에 필요한 공정 제어정보 생성 장치.

**청구항 5**

제1항에 있어서, 상기 드리프트 특징정보 생성부는,

제1 필터를 이용하여 상기 제1 특징 맵을 필터링함에 따라, 동일 공정과정에서의 상기 공정 파라미터 정보들간의 연관성을 고려한 제2 특징 맵을 생성하는 제2 특징 맵 생성부; 및

상기 제1 필터와는 다른 스케일의 제2 필터를 이용하여 상기 제1 특징 맵을 필터링함에 따라, 동일한 공정 파라미터 정보를 기준으로 연속된 공정 과정들간의 연관성을 고려한 제3 특징 맵을 생성하는 제3 특징 맵 생성부;를 더 포함하는 것을 특징으로 하는 기관 제조공정에 필요한 공정 제어정보 생성 장치.

**청구항 6**

제1항에 있어서,

상기 제2 공정 제어정보, 기관을 제조한 이전 공정에서의 공정 제어 히스토리 정보 및 상기 이전 공정에서 제조되었던 기관의 영역별 스케일 히스토리 정보들을 입력 받아 상기 입력부로 전달하는 업데이트부;를 더 포함하는 것을 특징으로 하는 기관 제조공정에 필요한 공정 제어정보 생성 장치.

**청구항 7**

제1항에 있어서,

상기 생성망은 상기 판별망으로부터 도출된 평가 결과를 고려하여, 상기 복수개의 레이어들 중 적어도 하나의 레이어의 필터 계수를 변경하는 것을 특징으로 하는 기관 제조공정에 필요한 공정 제어 정보 생성 장치.

**청구항 8**

제1항에 있어서,

상기 생성망 및 상기 판별망은 최소 제곱법을 기반으로 하는 손실 함수(least square loss function)를 이용하는 것을 특징으로 하는 기관 제조공정에 필요한 공정 제어 정보 생성 장치.

**청구항 9**

제1항에 있어서,

상기 생성망 및 상기 판별망은 베이지안 최적화 모델(Bayesian optimization)을 기반으로 변경하고자 하는 필터 계수들을 변경 설정하는 것을 특징으로 하는 기관 제조공정에 필요한 공정 제어 정보 생성 장치.

**청구항 10**

제조하고자 하는 기관에 속하는 영역들의 스케일과 관련된 목표 스케일 정보 및 상기 기관을 제조하기 위한 적어도 하나의 장비에 대한 공정 파라미터 정보를 입력 받는 단계;

상기 목표 스케일 정보 및 상기 공정 파라미터 정보를 기초로 복수개의 레이어들을 이용하여 상기 기관을 제조하기 위한 공정에 필요한 제1 공정 제어정보를 생성하는 단계; 및

상기 제1 공정 제어정보 및 트레이닝 데이터인 제2 공정 제어정보를 기초로 학습함에 따라 상기 제1 공정 제어정보에 대한 평가 결과를 도출하는 단계;를 포함하며,

상기 제1 공정 제어정보 생성 단계는,

상기 기관의 공정 제어와 관련되어 기 저장된 공정 제어 히스토리 정보를 이용하여 제1 특징 맵을 생성하고, 상기 제1 특징 맵을 서로 다른 스케일의 필터들을 각각 이용하여 컨볼루션 필터링 함에 따라 드리프트(drift) 특징정보를 생성하는 단계;를 포함하는 기관 제조공정에 필요한 공정 제어정보 생성 방법.

**청구항 11**

제10항에 있어서,

상기 스케일 정보는 상기 기관에 속하는 각 영역의 두께 정보이고,

상기 공정 파라미터 정보는 상기 장비의 누적 사용시간에 대한 정보이며,

상기 제1 공정 제어정보 및 제2 공정 제어정보는 단일 운용과정에서 소요되는 폴리싱 시간 및 폴리싱 압력값을 포함하는 것을 특징으로 하는 기관 제조공정에 필요한 공정 제어정보 생성 방법.

**청구항 12**

제10항에 있어서, 상기 제1 공정 제어정보 생성 단계는,

상기 드리프트 특징정보, 상기 스케일 정보들 및 상기 공정 파라미터 정보를 입력 받아 적어도 하나의 컨볼루션 레이어를 이용하여 컨볼루션 필터링하여 상기 제1 공정 제어정보에 대한 FC 특징정보를 생성하는 단계;를 더 포함하고,

상기 FC 특징정보를 디코딩하여 상기 제1 공정 제어정보를 출력하는 것을 특징으로 하는 기관 제조공정에 필요한 공정 제어정보 생성 방법.

**청구항 13**

폴리싱하고자 하는 기관에 속하는 영역들의 스케일과 관련된 목표 스케일정보를 입력 받아, 상기 기관을 폴리싱하기 위한 적어도 하나의 장비의 공정 파라미터 정보를 포함하는 현장 공정 제어를 위한 공정 제어 특징값을 생성하는 공정 제어정보 생성부; 및

상기 공정 제어 특징값을 기반으로 상기 장비의 공정조건을 제어하는 제어부;를 포함하며,

상기 공정 파라미터 정보는,

상기 장비의 압력 설정 정보 및 상기 장비의 폴리싱 시간과 관련한 시간 설정 정보를 포함하고,

상기 공정 제어정보 생성부는,

상기 기관의 공정 제어와 관련되어 기 저장된 공정 제어 히스토리 정보를 이용하여 제1 특징 맵을 생성하고, 상기 제1 특징 맵을 서로 다른 스케일의 필터들을 각각 이용하여 컨볼루션 필터링 함에 따라 드리프트(drift) 특징정보를 생성하는 드리프트 특징정보 생성부;를 포함하는 공정 제어 장치.

**청구항 14**

제13항에 있어서,

상기 제어부는 상기 제어된 장비의 공정조건 및 상기 장비의 누적 사용시간을 상기 공정 제어정보 생성부로 전달하고,

상기 목표 스케일 정보는 상기 기관에 속하는 각 영역의 목표 두께에 대한 정보인 것을 특징으로 하는 공정 제어 장치.

**청구항 15**

제13항에 있어서, 상기 공정 제어정보 생성부는,

상기 드리프트 특징정보, 상기 스케일 정보들 및 상기 공정 파라미터 정보를 입력 받아 적어도 하나의 컨볼루션 레이어를 이용하여 컨볼루션 필터링하여 상기 기관을 폴리싱하기 위한 공정에 필요한 제1 공정 제어정보에 대한 FC 특징정보를 생성하는 FC 특징정보 생성부;를 더 포함하고,

상기 FC 특징정보를 디코딩하여 상기 공정 제어정보를 생성하는 것을 특징으로 하는 공정 제어 장치.

**청구항 16**

제15항에 있어서, 상기 드리프트 특징정보 생성부는,

제1 필터를 이용하여 상기 제1 특징 맵을 필터링함에 따라, 동일 운용과정에서의 상기 공정 파라미터 정보들간의 연관성을 고려한 제2 특징 맵을 생성하는 제2 특징 맵 생성부; 및

상기 제1 필터와는 다른 스케일의 제2 필터를 이용하여 상기 제1 특징 맵을 필터링함에 따라, 동일한 공장 파라미터 정보를 기준으로 연속된 운용 과정들간의 연관성을 고려한 제3 특징 맵을 생성하는 제3 특징 맵 생성부;를 더 포함하는 것을 특징으로 하는 공정 제어 장치.

**청구항 17**

제10항 내지 제12항 중 어느 한 항에 따른 공정 제어정보 생성 방법을 컴퓨터에서 실행시키기 위하여 컴퓨터에서 판독 가능한 기록 매체에 저장된 컴퓨터 프로그램.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 공정 제어정보 생성 장치, 방법 및 이를 포함하는 공정 제어장치에 관한 것이다.

**배경 기술**

[0002] 반도체 제조 공정은 매우 복잡한 물리적 화학적 반응과 장비 상태에 따른 비선형적인 공정 프로세스의 변화

(process variation) 때문에 균일한 공정 결과를 지속적으로 얻기가 쉽지 않다. 정밀한 공정 제어를 위해 1990 년대에 제시된 run-to-run (R2R) control은 매 run (a single wafer 또는 wafer lot 등 가공 단위)마다 process variation를 추정하고 이를 입력에 반영하여 공정 결과가 타깃으로부터 벗어나지 않도록 한다.

[0003] 공정 프로세스의 변화의 유형에 해당하는 프로세스 드리프트(process drift)는 장비의 노후화에 따라 서서히 발생하는 프로세스 성능의 악화를 뜻하며, 또 다른 유형인 프로세스 시프트(shift)는 급격한 프로세스 성능의 변화, 예를 들어 장비의 소모품 교체로 인한 프로세스의 급격한 성능 개선을 뜻한다.

[0004] 전통적으로 R2R control은 선형 시스템 제어에 많이 쓰이는 exponential weighted moving average (EWMA) 방식과 double EWMA 방식을 사용해 왔다. EWMA 방식은 과거 run들의 제어 결과를 가중 평균해 process drift를 추정하는데, 계산 복잡성이 낮아 R2R controller에 우선적으로 고려되어 왔다. Double EWMA 방식은 공정 결과를 추정하는 EWMA filter와 process drift를 추정하는 EWMA filter의 복합 사용으로 process drift에 대처한다.

[0005] 반도체 제조에 필요한 많은 단위 공정 중 chemical mechanical planarization (CMP) 공정은 process drift가 뚜렷하게 발생하여 R2R control이 반드시 필요하다. 이밖에 3개의 EWMA filter를 사용한 R2R control 모델이 double EWMA 모델 대비 우수한 제어 성능을 보임을 실험을 통해 증명한 연구도 있다. 하지만 이런 EWMA 모델들은 비선형 process drift를 가진 시스템에 적용하기 어려워, drift가 급격하게 변하는 현상이 없을 시에만 사용하도록 권고하고 있다.

[0006] Neural network-based R2R (NNR2R) control 방법은 공정 인자들이 복잡한 상관관계를 가진 비선형 시스템의 제어에 매우 유용하다. Neural network는 시스템의 입출력 데이터만을 가지고 network 가중치 조정하기에 다양한 형태의 함수를 학습할 수 있어 비선형 시스템을 모델링하고 제어하는데 효과적이다. CMP 공정에서 NNR2R controller를 사용하여 wafer의 두께가 균일한 profile을 갖도록 입력을 제어하는 것이 가능함을 입증했다.

[0007] NNR2R controller는 2개의 neural network인 process model과 control model로 구성되며, 여기서 process model은 virtual metrology (VM)을 의미한다. VM 모델은 실제 계측에 필요한 비용과 시간을 줄여주거나, process drift를 신속하고 정확하게 감지할 수 있으며, 불량 wafer를 조기에 발견하여 최종 생산량을 향상시킨다. 이러한 이점으로 최근 기계 학습 방법을 활용하여 VM의 예측 능력을 극대화하기 위한 연구들이 지속되고 있다. VM 연구에는 process drift를 근사하기 위해 neural network를 결합한 VM 모델 구조를 제안한 연구와 에칭 공정의 trace data를 CNN에 사용한 VM 모델링 연구, 그리고 Gaussian process regression 기반 VM 모델을 통해 R2R 제어 성능을 향상시킨 연구 등이 있다. 현재까지 VM 모델에 대한 연구가 활발히 진행되고 있지만 R2R control model과 VM 모델을 통합할 때는 주의가 필요하다. R2R controller의 성능은 VM 모델의 예측 error와 control model의 error에 의해 결정되므로 단일 R2R control model에 비해 error가 높다. 특히 반도체 제조공정처럼 여러 가지 입력 요인이 intercorrelated되어 있고 비선형적인 process drift 현상을 가지는 경우에는 VM 모델의 예측 정확도가 높지 않으면 우수한 제어 성능을 보장하기가 어렵다.

**선행기술문헌**

**특허문헌**

[0008] (특허문헌 0001) 한국 공개 특허 제10-2019-0041908호 (공개)

**발명의 내용**

**해결하려는 과제**

[0009] 본 발명은 상기 전술한 종래의 문제점을 해결하기 위해 반도체 제조업에서 웨이퍼의 계측, R2R 제어, 이상 탐지 등을 아우르는 시스템으로 반도체 생산공정의 수율 향상과 공정의 신뢰성을 향상시키는, 공정 제어정보 생성 장치, 방법 및 이를 포함하는 공정 제어장치를 제공한다.

**과제의 해결 수단**

[0010] 상기 목적을 달성하기 위한 본 발명의 실시예에 따른 공정 제어정보 생성 장치는, 제조하고자 하는 기판에 속하는 영역들의 스케일과 관련된 목표 스케일 정보 및 상기 기판을 제조하기 위한 적어도 하나의 장비에 대한 공정 파라미터 정보를 입력 받는 입력부, 상기 목표 스케일 정보 및 상기 공정 파라미터 정보를 기초로 복수개의 레

이어들을 이용하여 상기 기관을 제조하기 위한 공정에 필요한 제1 공정 제어정보를 생성하는 생성망 및 상기 제1 공정 제어정보 및 트레이닝 데이터인 제2 공정 제어정보를 기초로 학습함에 따라 상기 제1 공정 제어정보에 대한 평가 결과를 도출하는 판별망을 포함할 수 있다.

[0011] 상술한 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 공정 제어정보 생성 방법은, 제조하고자 하는 기관에 속하는 영역들의 스케일과 관련된 목표 스케일 정보 및 상기 기관을 제조하기 위한 적어도 하나의 장비에 대한 공정 파라미터 정보를 입력 받는 단계, 상기 목표 스케일 정보 및 상기 공정 파라미터 정보를 기초로 복수개의 레이어들을 이용하여 상기 기관을 제조하기 위한 공정에 필요한 제1 공정 제어정보를 생성하는 단계 및 상기 제1 공정 제어정보 및 트레이닝 데이터인 제2 공정 제어정보를 기초로 학습함에 따라 상기 제1 공정 제어정보에 대한 평가 결과를 도출하는 단계를 포함할 수 있다.

[0012] 상술한 목적을 달성하기 위한 본 발명의 또 다른 실시예에 따른 공정 제어 장치는, 폴리싱하고자 하는 기관에 속하는 영역들의 스케일과 관련된 목표 스케일정보를 입력 받아, 상기 기관을 폴리싱하기 위한 적어도 하나의 장비의 압력 설정 정보 및 상기 장비의 폴리싱 시간과 관련한 시간 설정 정보를 포함하는 공정 제어정보를 생성하는 공정 제어정보 생성부 및 상기 공정 제어정보를 기반으로 상기 장비의 공정조건을 제어하는 제어부;를 포함하되, 상기 제어부는 상기 제어된 장비의 공정조건 및 상기 장비의 누적 사용시간을 상기 공정 제어정보 생성부로 전달할 수 있다.

[0013] 또한 본 발명은 상기한 방법에 따른 공정 제어정보 생성 방법을 실행시키는 컴퓨터로 판독 가능한 기록매체에 저장된 컴퓨터 프로그램을 제안한다.

**발명의 효과**

[0014] 본 발명의 실시예에 따른 공정 제어정보 생성 장치, 방법 및 이를 포함하는 공정 제어장치는 장비의 노후화에 의해 발생하는 프로세스 성능 악화에 따른 프로세스 변화를 정확하게 추정하여 웨이퍼를 효과적으로 폴리싱하기 위한 제어값들을 제공할 수 있는 효과가 있다

**도면의 간단한 설명**

- [0015] 도1은 일 실시예에 따른 기관(웨이퍼) 제조 공정에 사용되는 부품 및 장비를 도시한 도면이다.
- 도2는 본 발명의 일 실시예에 따른 공정 제어 장치의 구성을 개략적으로 도시한 블록도이다.
- 도3은 본 발명의 일 실시예에 따른 공정 제어정보를 출력하기 위한 학습과정을 설명하기 위해 도시한 개념도이다.
- 도4는 본 발명의 일 실시예에 따른 생성망의 구성을 개략적으로 도시한 블록도이다.
- 도5는 본 발명의 일 실시예에 따른 드리프트 특징정보 생성부를 설명하기 위해 도시한 도3의 확대도이다.
- 도6은 본 발명의 일 실시예에 따른 공정 제어 장치의 구성을 개략적으로 도시한 블록도이다.
- 도7은 성능 평가를 위한 지표를 도시한 도면이다.
- 도8은 도7에서의 두 비교 모델 별로 hyper-parameter를 B0로 최적화하여 예측 성능을 높이는 실험을 수행한 결과이다.
- 도9는 hyper-parameter 최적화를 위한 실험 결과를 나타낸 도면이다.
- 도10는 hyper-parameter 최적화를 위한 실험 결과를 나타낸 도면이다.
- 도11은 R2R 제어 알고리즘을 실험한 결과를 나타낸 도면이다.

**발명을 실시하기 위한 구체적인 내용**

[0016] 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계 없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.

[0017] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함" 한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다

른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록"등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

- [0018] 이하, 본 발명의 일 실시예를 첨부된 도면들을 참조하여 상세히 설명한다. 또한, 본 발명을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략할 수 있다.
- [0019] 반도체 사업에서는 단위 면적당 더 많은 반도체 소자를 만들기 위해 미세 회로 레이어를 웨이퍼(wafer)에 적층하는 구조가 필요하게 되었다. 하지만 미세 회로 패턴을 형성하는 lithography 공정과 필요한 회로 패턴을 제외한 나머지 레이어 부분을 제거하는 에칭(etching) 공정이 반복되면서 회로가 생성된 높은 부분과 나머지 낮은 부분의 표면 단차가 발생된다. 표면 단차는 후속 레이어의 회로 패턴을 형성하는 lithography 공정에 해상도에 치명적인 요소로 작용하여 반도체 소자의 불량률을 야기하게 된다. 따라서 적층 구조를 구현하기 위해서는 웨이퍼의 국부적인 평탄화 및 미세 회로 패턴에 필요한 광역적인 평탄화가 가능한 CMP 공정이 필요하게 되었다. CMP(chemical-mechanical polishing) 공정은 도1과 같은 로터리(rotary) 장비를 사용한다. 이 장비는 웨이퍼를 장착하는 캐리어(carrier)와 폴리싱 패드(polishing pad)를 장착하는 플레튼(platen), 패드(pad)를 청소하는 드레서(dresser), 그리고 연마 슬러리(slurry)가 공급되는 장치로 구성될 수 있다. CMP 공정은 연마 입자가 포함된 슬러리와 웨이퍼 표면의 소재 사이에서 발생하는 화학적 작용과 폴리싱 패드와 웨이퍼의 직접적인 접촉에 의한 기계적 작용을 통해 웨이퍼의 필름을 평탄화한다. CMP의 소모성 부품(폴리싱 패드, 리테이너 링(retainer ring), 멤브레인(membrane), 및 드레서)을 사용하는 대표적 공정이다. 소모성 부품은 수명에 따라 소재제거율(material removal rate)을 달리하게 만드는 드리프트(drift)를 발생시키고, 또한 장비의 소모성 부품을 교체하면 제거율이 개선되는 프로세스 시프트(process shift)도 발생한다. 이런 프로세스의 다양한 변화 때문에 CMP 공정은 제어하기 힘든 대표적인 비선형 시스템이다.
- [0020] CMP 공정의 품질은 웨이퍼에 남은 소재의 두께이다. Rotary CMP 장비는 웨이퍼의 영역별 두께를 제어하기 위해 웨이퍼에 맞닿아 있는 캐리어의 멤브레인이 각 영역별로 다른 압력을 가할 수 있도록 설계되어 있어 영역별로 소재의 두께를 관리할 수 있다. 도1에 도시된 Z1~Z7은 압력이 가해지는 웨이퍼의 영역들(zones)을 나타낸다.
- [0021] CMP 공정 제어를 위해서는 제어 변수들(control variables)과 부품 파라미터들(equipment parameters)이 필요하다. 제어 변수로는 플레튼(platen)과 캐리어(carrier)의 상대속도, 폴리싱 시간, 슬러리 유량, 드레서(dresser)의 속도, 그리고 웨이퍼(기판)의 영역별로 패드에 누르는 압력 등이 있지만, 실제 생산 공정에서 영역의 압력과 폴리싱 시간을 제외한 제어 변수들은 기존 사용 값에서 변화량이 크지 않도록 엄격하게 통제되어 있다. 부품 파라미터들은 소모품의 수명을 기록하며 프로세스 변화 모델링(process variation modeling)을 위한 정보로 활용된다.
- [0022] 이하에서는 전술한 바와 같은 비선형 시스템인 CMP 공정에 최적화하여 다양한 환경 변화에도 불구하고 최적의 웨이퍼(기판)의 영역별 원하는 두께를 구현할 수 있도록 하는 본 발명의 실시예에 따른 공정 제어 정보 생성 장치, 공정 제어 장치 및 방법의 구성을 관련된 도면을 참조하여 상세히 설명한다.
- [0023] 본 발명의 실시예에 따른 공정 제어 정보 생성장치는 매 공정(run)마다 웨이퍼(기판)의 각 영역별로 타겟 두께를 얻기 위해 부품 파라미터들을 고려하여 영역의 압력과 폴리싱 시간을 결정하는 제어 모델을 수립하는 것을 목적으로 한다.
- [0024] 본 발명의 실시예에 따른 공정 제어 장치는 입력부, 공정 제어 정보 생성부, 제어부, 및 업데이트부를 포함하여 구성될 수 있다. 상기 공정 제어 장치는 이미 이전에 기계학습이 완료됨에 따라 내부의 각 레이어의 필터 계수들이 설정된 상태로 실제 공정에 적용되었을 때 이용되는 장치인 것이다. 이와 같은 공정 제어 장치는 공정에 적용되기 전 학습 과정에 적용되는 공정 제어정보 생성 장치에 대하여 설명한 후 자세하게 설명하도록 한다.
- [0025] 도2는 본 발명의 일 실시예에 따른 공정 제어정보 생성 장치의 구성을 개략적으로 도시한 블록도이다. 도2를 참조하면, 본 발명의 실시예에 따른 공정 제어정보 생성 장치는, 입력부(100), 학습부(200), 제어부(300) 및 업데이트부(400)를 포함하여 구성될 수 있다.
- [0026] 본 발명의 실시예에 따른 입력부(100)는 제조하고자 하는 기판에 속하는 영역들의 스케일과 관련된 목표 스케일 정보 및 기판(이하, 웨이퍼)을 제조하기 위한 적어도 하나의 부품에 대한 공정 파라미터 정보를 입력 받을 수 있다.
- [0027] 여기서, 목표 스케일 정보란 웨이퍼를 폴리싱 하는데 있어서 달성하고자 하는 웨이퍼의 영역별 목표 두께 값을

의미하는 것으로서 사용자로부터 입력받을 수 있고, 공정 파라미터 정보는 도1에 도시된 바와 같은 웨이퍼를 폴리싱하기 위해 사용되는 부품들이 웨이퍼를 폴리싱하는데 사용된 누적 시간을 의미할 수 있다.

[0028] 일 실시예에 따른 입력부(100)는 도2에 도시된 바와 같이, 별도로 마련되어 목표 두께 값 및 부품의 누적 사용 시간에 대한 정보를 입력 받아, 생성망(210) 및 판별망(220)으로 전달하도록 구현될 수도 있고, 입력부(100) 없이 바로 입력 정보들(목표 두께 값, 부품 사용 누적 시간 등) 각각이 바로 생성망(210) 또는 판별망(220)으로 입력되어 학습될 수도 있다.

[0029] 그리고, 본 발명의 실시예에 따른 학습부(200)는 상술한 기계 학습이 완료된 이후 실제 공정에 적용시키기 위한 공정 제어 장치의 공정 제어정보 생성부의 각 레이어의 필터계수를 설정하기 위해서 기계 학습하는 구성으로, 본 발명의 학습부(200)는 생성망(210) 및 판별망(220)으로 구성될 수 있다.

[0030] 생성망(210)은 입력부(100)에 입력된 스케일 정보, 및 공정 파라미터 정보를 기초로 복수개의 레이어들을 이용하여 웨이퍼를 제조하기 위한 공정에 필요한 제1 공정제어 정보를 생성할 수 있다.

[0031] 판별망(220)은 제1 공정 제어정보 및 트레이닝 데이터로서 실제 기관 제조공정에 적용되었던 제2 공정 제어정보를 기초로 생성망(210)에 의해 생성된 제1 공정 제어정보에 대한 평가 결과를 도출할 수 있다.

[0032] 도2와 같은 구조의 공정 제어정보 생성 장치는 매 공정(run) 과정이 완료될 때마다 업데이트부(400)를 이용하여 최근 가공된 공정 데이터(run data)를 입력부 및 학습부로 전달함에 따라 학습할 수 있다.

[0033] 여기서, 본 발명의 학습부(200)는 생성망(210)과 생성망 학습을 경쟁을 통해 도와주는 판별망(220)으로 구성된 LSGAN(Least Squares Generative Adversarial Network)이다. 위에서 설명한 종래의 NNR2R(neural network-based run to run control)이 VM(virtual metrology) 모델을 이용하여 제어 값을 찾는 간접 제어 방식인 반면에, 본 발명의 생성망(210)은 제어-출력간 관계를 판별망(220)과의 경쟁을 통해 학습하여 R2R 제어에 직접 사용한다. 따라서, 학습된 생성망(210)은 목표 출력값(목표 스케일 정보)을 입력 받으면, 달성하기 위한 공정 제어 정보(제어 값)을 출력할 수 있다.

[0034] 본 발명의 생성망(200)은 아래 <수학식1>과 같은 제어 함수를 이용하여 공정 제어정보를 생성할 수 있다.

**수학식 1**

[0035] 
$$u_k = f(y^{\text{target}}, y_{k-1}, y_{k-2}, \dots, y_{k-n}, u_{k-1}, u_{k-2}, \dots, u_{k-n})$$

[0036] 여기서,  $u_k$ 는 생성망(210)으로부터 출력되는 출력 제어 값이고,  $y$ 는 출력 벡터(output vector),  $u$ 는 제어 벡터(control vector),  $n$ 은 생성망(210)에서 참조되는 가공된 공정(run)의 수를 의미한다. 생성망(210)은 업데이트부(400)로부터 가장 최근의 공정 제어 히스토리 정보(제어값) 그리고 출력 데이터와 폴리싱 공정에 이용되는 부품들의 누적 사용시간을 제공 받아, 최종 폴리싱되는 웨이퍼의 영역별 출력 두께값( $y_k$ )과 사용자로부터 입력부(100)로 입력된 목표 두께값( $y^{\text{target}}$ )이 같아지도록 하는 공정 제어정보( $u_k$ )를 도출할 수 있다.

[0037] 이때, 공정 제어 히스토리 정보는 공정의 시퀀스 정보, 각각의 장비를 제어하기 위한 제어 파라미터 정보를 포함할 수 있다. 또한, 공정 제어 히스토리 정보는 상기 시퀀스에서의 목표 스케일 정보를 더 포함할 수 있다. 여기에서 제어 파라미터 정보는, 예를 들어, 과거 복수번의 공정들 각각에서의 제어값, 즉 과거 공정에서 설정되었던 부품의 압력값 및 폴리싱 시간, 생성망의 각 레이블의 필터 계수(가중치), 최종 폴리싱된 영역별 두께값들을 포함하는 웨이퍼 폴리싱 공정에서의 모든 파라미터 정보들을 의미할 수 있다.

[0038] 상기와 같은 공정 제어정보를 이용하여 생성망(210)이 제1 공정 제어정보를 생성하기 위한 제어 함수는 아래 <수학식2>와 같이 정의될 수 있다.

**수학식 2**

[0039] 
$$\hat{u}_k = G(y^{\text{target}}, y_{k-1}, y_{k-2}, \dots, y_{k-n}, u_{k-1}, u_{k-2}, \dots, u_{k-n}, \emptyset, W_{\text{drift}}, W_{R2R})$$

- [0040] 여기서,  $G(\cdot)$ 는 생성망(210)에서 이용하는 함수이고,  $\hat{u}_k$ 는 생성망(210)이 생성하는 공정 제어정보이며,  $\Phi$ 는 공정 파라미터 정보인 부품의 누적 사용시간이고,  $W_{drift}$  및  $W_{R2R}$ 은 생성망에서의 학습 가중치(레이어들의 필터계수)이다.
- [0041] CMP 공정의 프로세스 드리프트(process drift)에 대처하기 위해서는 본 발명의 공정 제어정보 생성 장치에서의 매 공정(run)이 완료될 때마다 최근 공정의 공정 파라미터 정보를 기반으로 생성망(210)의 가중치( $W_{drift}$ )를 갱신할 수 있다. 이어서, 생성망(210)과 판별망(220)의 경쟁적 학습을 통해 목표 스케일 정보(목표 두께 값)을 달성하기 위한 가중치( $W_{R2R}$ )를 갱신하고 학습이 완료되면 목표 두께 값( $y^{target}$ )을 달성하기 위한 공정 제어정보( $u_k$ )를 출력할 수 있다.
- [0042] 종래의 신경망인 GAN은 [0, 1] 범위 내에서 정의된 임의의 latent 벡터(z)에 대응되는 출력 값이 실제 출력 값과 같아지도록 학습이 진행된다. 반면에 본 발명의 공정 제어정보 생성 장치에서의 현재 공정의 제어 변수들은 이전 공정들에서의 제어변수들과 시간적으로 종속적인 관계에 있다. 따라서, 본 발명의 공정 제어정보 생성 장치는 종래의 GAN과 같이 latent 벡터(z)에 대한 제어 변수 값들(출력 값)을 학습하지 않고, 이전 공정들에서의 공정 제어 변수들에 대한 특징정보를 기초로 하여 현재 공정에서의 공정 제어변수 값을 예측할 수 있도록 한다.
- [0043] 도3은 본 발명의 일 실시예에 따른 공정 제어정보를 출력하기 위한 학습과정을 설명하기 위해 도시한 개념도이고, 도4는 본 발명의 일 실시예에 따른 생성망(210)의 구성을 개략적으로 도시한 블록도이다.
- [0044] 도3 및 도4에 도시된 바와 같이, 본 발명의 생성망(210)은 드리프트 특징정보 생성부(211), FC 특징정보 생성부(212), 및 출력부(213)를 더 포함하여 구성될 수 있다. FC 특징정보 생성부(212)에 현재 공정에서의 부품 누적 사용시간에 대한 정보(11)와 웨이퍼의 영역별 목표 두께값(12)을 입력하면, 학습된 FC 특징정보 생성부(212)는 공정 제어정보인 영역별 폴리싱 시간과 폴리싱 압력값을 출력한다. 프로세스의 변화들을 공정 제어에 반영하기 위해 생성망(210)은 FC 특징정보 생성부(212) 앞에 드리프트 특징정보 생성부(211)를 구성한다. 여기서, 본 발명의 드리프트 특징정보 생성부(211)는 CNN(Convolution Neural Network)일 수 있다. 드리프트 특징정보 생성부(211)는 매 공정(run)이 진행됨에 따라 슬라이딩 윈도우(sliding window)를 이용하여 최근 처리되었던 과거 공정들(runs)의 데이터를 업데이트부(400)로부터 전달 받아 공정 제어 히스토리 정보를 업데이트 할 수 있다. 그 후, FC 특징정보 생성부(212)는 업데이트된 공정 제어 히스토리 정보를 기반으로 드리프트 특징정보 생성부(211)로부터 출력되는 드리프트 특징 정보를 입력 받아, 판별망(220)과 경쟁적 학습을 진행하고, 이와 같은 학습이 완료되면 폴리싱을 위한 부품을 제어하기 위한 공정 제어정보(제어 값)를 출력할 수 있다.
- [0045] 상술한 내용을 도3 내지 도5를 참조하여 각 구성에 대한 동작을 순서에 따라 보다 상세하게 설명한다. 도5는 본 발명의 일 실시예에 따른 드리프트 특징정보 생성부를 설명하기 위해 도시한 도3의 확대도이다. 도5를 참조하면, 본 발명의 실시예에 따른 드리프트 특징 정보 생성부(211)는 인코딩부(510), 제2 특징 맵 생성부(520) 및 제3 특징 맵 생성부(530)를 더 포함하여 구성될 수 있다.
- [0046] 본 발명의 입력부(100)는 공정 파라미터 정보인 부품 누적 사용시간에 대한 정보(11), 웨이퍼의 각 영역에 대한 목표 두께 값(12), 및 공정 제어 히스토리 정보와 스케일 히스토리 정보(13)를 입력 받는다.
- [0047] 그리고, 본 발명의 실시예에 따른 드리프트 특징정보 생성부(211)의 인코딩부(510)는 이전에 제조된 웨이퍼들 각각의 시퀀스(sequence) 정보, 웨이퍼들을 제조할 때 사용되었던 부품에 대한 공정 파라미터 정보, 이전에 제조된 웨이퍼들 각각을 제조할 때 소요되었던 폴리싱 시간, 및 폴리싱 압력값을 포함하는 공정 제어 히스토리 정보(13)를 이용하여 제1 특징 맵(511)을 생성할 수 있다.
- [0048] 일 실시예에 따른 제1 특징 맵(511)은 웨이퍼별 각 시퀀스 정보(세로 축)에 따른 상기 공정 파라미터 정보, 폴리싱 시간, 폴리싱 압력, 및 폴리싱된 웨이퍼의 영역별 출력 두께 값(가로 축)들을 표현한 것으로서, 0 내지 1 사이의 특징 값으로 표현된 특징 맵일 수 있다.
- [0049] 그리고, 본 발명의 실시예에 따른 제2 특징 맵 생성부(520)는 제1 필터(521)를 이용하여 제1 특징 맵(511)(컨볼루션 레이어)을 필터링함에 따라, 동일 공정과정에서의 공정 파라미터 정보들간의 연관성을 고려한 제2 특징 맵을 생성할 수 있다.
- [0050] 일 예로, 제2 특징 맵 생성부(520)는 1 X 2 크기의 제1 필터(521)를 이용하여 제1 특징 맵(511)을 컨볼루션 필터링함에 따라 웨이퍼들 간의 공정 변화와 관련된 제2 특징 맵을 생성할 수 있다.

[0051] 또한, 본 발명의 실시예에 따른 제3 특징 맵 생성부(530)는 제2 필터(531)(컨볼루션 레이어)를 이용하여 제1 특징 맵(511)을 컨볼루션 필터링함에 따라, 동일한 공정 파라미터 정보를 기준으로 연속된 공정 과정들간의 연관성을 고려한 제3 특징 맵을 생성할 수 있다.

[0052] 일 예로, 제3 특징 맵 생성부(530)는 2 X 1 크기의 제2 필터(531)를 이용하여 제1 특징 맵(511)을 컨볼루션 필터링함에 따라 공정 제어 히스토리 정보들 간의 공정 변화와 관련된 제3 특징 맵들을 생성할 수 있다.

[0053] CNN인 드리프트 특징정보 생성부(211)는 복수개의 컨볼루션 레이어(convolution layer)들과 풀링 레이어(pooling layer)들을 적층한 구조이다. 드리프트 특징정보 생성부(211)는 제1 필터 및 제2 필터를 이용하여, 컨볼루션 레이어인 제2 특징 맵 및 제3 특징 맵을 생성하며, 풀링 레이어들은 제2 특징 맵 및 제3 특징 맵에서 대표적인 특징값을 선택하여 각각의 특징 맵의 크기를 줄인다. 이러한 방법으로 드리프트 특징정보 생성부(211)는 process drift를 고려한 드리프트 특징정보를 생성하여, FC 특징정보 생성부(212)로 전달한다. 이렇게 생성된 드리프트 특징정보는 FC 특징정보 생성부(212)의 FC 레이어(fully connected layer)에 전달되어 이미지 인식이나 분류에 사용된다.

[0054] 즉, 도3에 도시된 바와 같은 최근 n개의 과거에 폴리싱된 공정 제어정보들은 2차원 공정 이미지로 생각할 수 있다. 제1 필터(521)는 동일 공정에서의 변수들 간의 연관성을 고려하여 특징 맵을 생성하는 것이라면, 제2 필터(531)는 연속된 공정들간의 상관성을 고려한 특징 맵을 생성하는 것일 수 있다.

[0055] 제2 특징 맵과 제3 특징 맵 각각은 아래 <수학식3>에 의해 생성될 수 있다.

**수학식 3**

$$v_{ij}^1 = \alpha \left( \sum_{q=0}^1 \tilde{x}_{si,sj+q} w_{0,q} + b \right), 0 \leq i \leq \frac{H-1}{s}, 0 \leq j \leq \frac{W-2}{s},$$

$$v_{ij}^2 = \alpha \left( \sum_{q=0}^1 \tilde{x}_{si+p,sj} w_{p,0} + b \right), 0 \leq i \leq \frac{H-2}{s}, 0 \leq j \leq \frac{W-1}{s}$$

[0056]

[0057] 여기서,  $v_{ij}^1$ 는 1X2 크기의 제1 필터를 사용한 제2 특징 맵의 노드(i, j)이며,  $v_{ij}^2$ 는 2X1 크기의 제2 필터를 사용한 제3 특징 맵의 노드(i, j)이다.  $\alpha$ 는 활성화 함수이고, H와 W는 입력 데이터의 열과 행의 크기이며, p와 q는 각 필터의 행과 열이고, s는 각 필터의 이동 간격인 stride length를 의미한다. 또한,  $\tilde{x}_{si,sj+q}$ 는 입력 데이터의 (si+p, sj+q) 좌표를 의미하고,  $w_{p,q}$ 와 b는 각 필터의 p와 q에 위치하는 가중치(weight)와 바이어스(bias)를 의미한다. 또한, 위 <수학식2>에서의  $w_{drift}$ 는 컨볼루션 레이어들의 가중치들을 의미한다.

[0058] 본 발명의 실시예에 따른 FC 특징정보 생성부(212)는 도3에 도시된 바와 같이, 드리프트 특징정보 생성부(211)로부터 출력된 드리프트 특징정보와 입력부(100)로부터 공정 파라미터 정보인 부품 누적 사용시간에 대한 정보 및 영역별 목표 두께 값에 대한 정보를 입력 받는다. 입력된 정보들 각각에 대하여 인코딩하여 특징정보(특징 벡터 또는 특징 행렬)로 변환하고, 변환된 특징 정보들을 적어도 컨볼루션 레이어들을 이용하여 컨볼루션 필터링함에 따라 FC 특징정보를 생성할 수 있다. 예컨대, FC 특징정보는 특징 벡터 또는 특징 행렬로 이루어진 특징 행렬로 생성될 수 있고, FC 특징정보는 생성된 FC 특징정보를 디코딩하여 부품들 각각에 적용 제어할 폴리싱 시간과 영역별 압력값을 포함하는 제1 공정 제어정보(15)를 출력할 수 있다.

[0059] 그리고, 이렇게 FC 특징정보 생성부(212)로부터 출력된 제1 공정 제어정보(15)는 판별망(220)으로 전달된다.

[0060] 도3에 도시된 바와 같은, 판별망(220)은 상기 제1 공정 제어정보(15) 및 트레이닝 정보로서 실제 폴리싱 시간 및 영역별 압력값을 포함하는 실제 웨이퍼 제조공정에 적용되었던 제2 공정 제어정보(14)를 입력 받을 수 있다.

[0061] 제2 공정 제어정보(14)는 도면에는 별도로 도시하지 않았으나, 본 발명의 공정 제어정보 생성 장치 내에 포함된 DB에 저장될 수 있고, DB는 생성망(210)에서 제1 공정 제어정보(15)를 출력하면 저장하고 있던 제2 공정 제어정보(14)를 판별망(220)으로 전달할 수 있다.

[0062] 제2 공정 제어정보(14)는 과거 웨이퍼들 각각을 폴리싱 할 때의 부품의 사용시간, 부품의 압력 값 등을 포함하는 실제 정보로서, 업데이트부(400)에 의해 전달 받은 입력부(100)가 DB로 저장되었던 정보일 수 있다.

[0063] 웨이퍼의 공정이 완료될 때마다 해당 공정의 제어정보(제어값)과 그에 따라 폴리싱된 웨이퍼의 영역별 두께 값(출력값) 그리고, 부품의 사용시간은 업데이트부(400)에 저장된다.

[0064] 이에 따라, 판별망(220)은 제1 공정 제어정보(15) 및 제2 공정 제어정보(14)를 인코딩하여 특징정보로 변환하고, 변환된 특징정보를 컨볼루션 레이어들을 이용하여 상기 제1 공정 제어정보는 0이 되도록, 제2 공정 제어정보는 1이 되도록 학습함에 따라 각 공정 제어정보에 판별을 수행할 수 있다.

[0065] 즉, 판별망(220)은 각 공정 제어정보를 입력 받아, 내부의 컨볼루션 레이어들을 이용하여 각각의 공정 제어정보에 대하여 판별함으로써 0 내지 1 사이의 값을 도출할 수 있다. 생성망(210)은 판별망(220)으로부터 제1 공정 제어정보(15)에 대한 출력값이 1에 수렴할 수 있도록 이후 학습단계에서는 내부 레이어들의 필터 계수(가중치)들을 변경하여 설정할 수 있다.

[0066] 상술한 바와 같은, 생성망(210)과 판별망(220)으로 구성된 본 발명의 학습부(200)는 LSGANs의 최소 제곱법을 기반으로 하는 손실 함수(least square loss function)를 사용하여 역전파(back propagation)를 통해 학습되며, 학습에 필요한 손실 함수는 아래 <수학식4>와 같다.

**수학식 4**

$$L_D^{LRM} = E_{\mathbf{u}_k \sim p_d} [(D(\mathbf{u}_k) - 1)^2] + E_{\hat{\mathbf{u}}_k \sim p_g} [D(\hat{\mathbf{u}}_k)^2],$$

$$L_G^{LRM} = E_{\hat{\mathbf{u}}_k \sim p_g} [(D(\hat{\mathbf{u}}_k) - 1)^2].$$

[0067]

[0068] 여기서,  $L_D^{LRM}$  과  $L_G^{LRM}$  은 판별망(220)과 생성망(210)의 손실(loss)이며,  $p_d$ 와  $p_g$ 는 제2 공정 제어정보( $\mathbf{u}_k$ )와 생성망(210)에 의해 생성된 제1 공정 제어정보( $\hat{\mathbf{u}}_k$ )의 분포이다. 판별망(220)은 제2 공정 제어정보( $\mathbf{u}_k$ )에 대해  $D(\mathbf{u}_k)=1$ , 제1 공정 제어정보  $D(\hat{\mathbf{u}}_k)=0$ 이 되도록 학습한다. 반면, 생성망(210)은 제2 공정 제어정보( $\mathbf{u}_k$ )와 아주 유사한 정보를 생성하여  $D(\hat{\mathbf{u}}_k)=1$ 이 되도록 학습한다. 궁극적으로 학습이 성공적으로 완료되면 분포  $p_g$ 는  $p_d$ 와 같아지게 된다.

[0069] 본 발명의 공정 제어정보 생성장치 LSGANs의 학습방식을 사용하는 이유는 생성망(210)이 학습한 다차원 제어 공간에서의 분포  $p_g$ 가 실제 분포인  $p_d$ 와 같아지도록 하는 것이 제어 성능을 높이기 때문이다.

[0070] 이론적으로 생성망(210)과 판별망(220)의 손실 함수는 학습을 통해 수렴이 가능하지만, 실제로는 신경망의 구조적 복잡성 때문에 수렴이 보장되지 않는다. 이 문제를 해결하기 위해 본 발명의 학습부(200)는 베이지안 최적화 모델(Bayesian optimization)을 기반으로 필터 계수들(hyper-parameters)을 조정하여 모델의 구조를 다르게 함으로써 학습이 수렴하게 될 수 있다. 본 발명의 학습부(200)는 베이지안 최적화 모델(BO)을 기반으로 하기 때문에 반복 검색을 통해 매 학습(iteration)마다 솔루션을 지능적으로 샘플링하기 때문에 빠른 시간 내에 우수한 모델 성능을 보장하는 필터 계수(가중치)들을 찾을 수 있다.

[0071] 아래 <수학식5>는 본 발명의 학습부(200)는 베이지안 최적화 모델(BO)을 기반으로 필터 계수들(hyper-parameter)를 조정하는 알고리즘에 대한 것이다.

**수학식 5**

$$t^* = \arg \min_{t \in T} c(t)$$

[0072]

[0073] 여기서,  $t$ 는 hyper-parameter solution,  $T$ 는 solution space,  $c(t)$ 는 학습 완료 후의 비용 함수(cost

function)을 의미한다. 위 <수학식5>의  $t^*$ 는 모델의 성능지표를 오차(error)라 하면, 오차가 가장 낮은 모델을 구성하는 hyper-parameter solution으로 정의될 수 있다.

[0074] 본 발명의 학습부(200)는 베이지안 최적화 모델(BO)을 적용하기 위해 아래 <수학식6>과 같은 네트워크 함수  $c(t)_{LRC}$ 를 정의한다.

**수학식 6**

$$c(t)_{LRC} = \frac{\sum_{i=1}^m \{ (d_{loss_i}(t) + g_{loss_i}(t) * |d_{loss_i}(t) - g_{loss_i}(t)|) \}}{m - 1}$$

[0076] 여기서, m은 LSGANs의 학습에 소요되는 에폭(epochs) 수를 의미하고, 각 에폭에서 판별망(220)과 생성망(210)이 배치(batch) 수만큼 번갈아 학습된다.  $d_{loss_i}(t)$ 와  $g_{loss_i}(t)$ 는 에폭(i)에서 판별망(220)과 생성망(210)의 손실을 뜻하며, l은 LSGANs의 학습이 안정되는 에폭 시점을 의미한다. LSGANs을 학습시킬 때 초기 에폭에서는 d\_loss와 g\_loss의 진동이 심하기 때문에  $c(t)_{LRC}$  값의 오류를 줄이기 위해 학습이 안정화되는 l 시점 이후의 에폭을 선택한다. 본 발명의 학습부의 전체 손실을 줄이고 d\_loss와 g\_loss의 적절한 균형을 맞추기 위해 d\_loss와 g\_loss의 합과 d\_loss와 g\_loss의 절대 차이를 곱하고 에폭 수인 m-1을 나눈 평균값을  $c(t)_{LRC}$ 로 정의한다.  $c(t)_{LRC}$ 의 값이 작을수록 생성망(210)이나 판별망(220)의 손실차이가 줄어든다는 의미로 생성망과 판별망의 어느 한쪽으로 치우치지 않는 상태에 도달한다고 볼 수 있다.

[0077] Solution(t)에 대해  $c(t)_{LRC}$ 는 학습이 완료되어야만 관측이 가능하므로 비용이 많이 든다. 따라서, solution space에서 효율적인 샘플링이 필요하다. BO는  $c(t)_{LRC}$ 가 Gaussian process prior를 가진다고 가정한다. 초기에 BO는 하나의 solution을 임의로 선택하여 학습부(200)를 구성하고, 트레이닝 정보(14)를 통해  $c(t)_{LRC}$ 를 평가한다. 관측된 데이터를 가지고 있는 BO는 포스테리어(posterior)를 계산하여  $c(t)_{LRC}$ 를 재추정하며, 업데이트된  $c(t)_{LRC}$ 에 기반한 획득 함수(acquisition function)을 통해 solution space에서 샘플링할 다음 지점을 선택한다. 이후 학습 종료의 조건이 만족될 때까지 샘플링을 반복하여 가장 좋은 solution( $t^*$ )를 찾는다.

[0078] 다시 도2를 참조하면, 이에 따라 본 발명의 실시예에 따른 제어부(300)는 생성망(210)으로부터 출력된 제1 공정 제어정보( $u^k$ )를 전달 받아, 웨이퍼를 폴리싱하기 위한 부품들 각각에 제어 정보(폴리싱 시간 및 폴리싱을 위한 압력값)( $y^k$ )을 입력하고, 이때 제어부(300)는 부품들로 전달한 제어 정보( $y^k$ )을 업데이트부(400)에도 전달할 수 있다. 즉, 웨이퍼의 공정이 완료될 때마다 해당 공정의 제어정보(제어값)가 업데이트부(400)에 저장될 수 있고, 더불어 상기 제어정보를 기반으로 폴리싱된 웨이퍼의 영역별 두께 값(출력값)과 부품의 사용시간 까지 업데이트부(400)에 저장될 수 있다.

[0079] 이에 따라, 본 발명의 실시예에 따른 업데이트부(400)는 저장된 제어정보(run data)를 학습부(200)의 이후 학습을 위하여 입력부(100)로 전달함으로써 업데이트 할 수 있다.

[0080] 다음으로, 도2 내지 도5를 참조하여 설명한 공정 제어정보 생성 장치를 통해 학습된 공정 제어정보 생성부를 이용하여 현장 웨이퍼 폴리싱 공정에서 실시간으로 공정을 제어하는 공정 제어 장치에 대하여 설명한다. 도6은 본 발명의 일 실시예에 따른 공정 제어 장치의 구성을 개략적으로 도시한 블록도이다.

[0081] 도6에 도시된 바와 같은 본 발명의 실시예에 따른 공정 제어 장치(600)는, 입력부(610), 공정 제어정보 생성부(620), 제어부(630), 및 업데이트부(640)를 포함하여 구성될 수 있다. 공정 제어 장치(600)의 입력부(610), 제어부(630), 및 업데이트부(640)는 도2에 도시된 입력부(100), 제어부(300), 및 업데이트부(400)의 기술적 동작과 실질적으로 유사하므로, 아래에서는 공정 제어정보 생성부(620), 및 제어부(630)를 위주로 설명한다.

[0082] 본 발명의 실시예에 따른 공정 제어정보 생성부(620)는, 폴리싱하고자 하는 웨이퍼에 속하는 영역들의 스케일과 관련된 목표 스케일 정보를 입력 받아, 웨이퍼를 폴리싱하기 위한 적어도 하나의 장비의 압력 설정정보 및 장비의 폴리싱 시간과 관련한 시간설정 정보를 포함하는 현장 공정 제어를 위한 공정 제어 특징값을 생성할 수 있다.

- [0083] 또한, 제어부(630)는, 공정 제어정보 생성부(620)로부터 생성된 공정 제어 특징값을 기반으로 장비의 공정조건을 제어할 수 있다.
- [0084] 공정 제어정보 생성부(620)는, 본 발명의 공정 제어정보 생성 장치의 학습과정을 통해 최적화된 가중치로 정해진 레이어들의 필터 계수들로 설정되어, 실시간 입력값들을 입력 받아 공정 제어 특징값을 생성함으로써, 제어부(630)는 생성된 공정 제어 특징값을 토대로 폴리싱을 위한 압력 값과 폴리싱 시간으로 장비들을 제어할 수 있다.
- [0085] 여기에서, 상기 공정 제어정보 생성부(620)로 입력되는 입력값들은 위에서 설명한 공정 제어정보 생성 장치의 생성망에 입력되는 입력 정보(값)들과 동일할 수 있다.
- [0086] 또한, 본 발명의 공정 제어정보 생성부(620)는 공정 제어 정보인 폴리싱을 위한 압력 값 및 폴리싱 시간으로 디코딩되기 전의 특징 벡터 또는 특징 행렬 형태인 공정 제어 특징값을 생성하여 출력하여 제어부(630)에서 상기 공정 제어 특징값을 디코딩함에 따라 압력 값과 폴리싱 시간을 획득하여, 획득된 압력 값과 폴리싱 시간으로 장비를 제어하도록 구현될 수 있고, 또 다른 예로, 상기 공정 제어 특징값을 생성하여 공정 제어 특징값을 디코딩하여 상기 압력 값과 폴리싱 시간을 제어부로 출력 전달하도록 구현될 수도 있다.
- [0087] R2R 제어기(controller)의 성능은 제어 오차에 기반해 측정하는데 여기서, 제어 오차는 각 공정(run)에 대해서 학습부의 출력과 실제 출력의 절대 차이를 의미한다. 본 발명에서는 각 제어 변수들별로 (SD)와 mean absolute error(MAE)를 계산하여 도7과 같이 모델들의 성능을 평가하였다. 도7에서 p는 제어 변수들(control variable), n은 테스트 데이터 크기, MAE는 학습된 모델이 실제 값을 얼마나 정확히 예측하는지를 나타내는 지표이며, 는 학습된 모델의 예측 정확도와 변동성을 함께 가늠할 수 있는 지표이다. 두 지표 모두 낮은 값을 기록할수록 제어 모델의 성능이 높다는 것을 나타내고 있다.
- [0088] 도8은 도7에서의 두 비교 모델 별로 hyper-parameter를 BO로 최적화하여 예측 성능을 높이는 실험을 수행한 결과이다. NNR2R(A)는 두 개의 신경망을 사용하고, 하나는 VM 모델로 활용하며, 다른 하나는 R2R controller로 사용한다. NNR2R(A)은 매 공정이 끝난 후 슬라이딩 윈도우에 의해 선택된 최신 공정 데이터를 이용해 VM 모델을 학습하고 학습된 VM 모델에 기반해 R2R controller를 학습한다. 따라서 NNR2R(A)의 R2R controller의 성능은 VM 모델의 성능에 크게 좌우된다. 본 발명에서는 VM 모델의 성능을 개선하기 위해 학습부에 적용한 프로세스 변화의 특징 추출을 위한 CNN을 VM 모델 에도 적용하였다.
- [0089] Hyper-parameter 최적화에서 BO의 연속 횟수는 100회, NNR2R(A)의 네트워크 함수는 
$$c(t)_{NNR2R} = \frac{\sum_{i=1}^m (VM_{loss_i}(t) + R2R_{loss_i}(t))}{m-1}$$
 로 설정하였다. 여기서 m은 학습 에폭의 수를 말하며 20회를 설정했고, l은 NNR2R의 학습이 안정화되는 에폭 시점을 의미하며 5회를 설정했다.  $VM_{loss_i}(t)$ 와  $R2R_{loss_i}(t)$ 는 epoch i에서 hyper-parameters solution(t)을 사용한 VM과 R2R controller의 각각 mean squared error를 뜻한다. 본 발명의 학습부(B)의 경우, NNR2R(A)과 동일하게 BO 연속 횟수는 100회, m=20, l=5로 설정했다. 도8은 최적화 실험 결과를 보여준다. 도8과 같은 실시예에서 학습주기(learning rate)는 0.000001부터 0.01사이이며, CNN 특징 맵의 개수는 4개에서 300개 사이이고, 각 FC 레이어(fully connected layer)의 은닉 노드의 개수는 10개에서 400개 사이이며, 풀링 레이어의 타입은 average pooling과 max pooling이다.
- [0090] 도9은 학습부의 hyper-parameters를 랜덤으로 찾는 방식의 알고리즘(a)과 BO 방식의 알고리즘(b)으로 최적화한 후의 각 학습 에폭별 생성망의 손실과 판별망의 손실을 보여주는 도면이다. 도9의 (a)는 생성망과 판별망의 손실이 학습 도중에 지속적으로 진도하며 수렴하지 않는 현상을 보여준다. 이러한 현상은 GANs의 학습과정에서 나타나는 대표적인 문제점으로 판별망과 생성망이 번갈아 가며 학습하므로 서로를 속고 속이는 과정에서 전역 해로 수렴할 수 없는 상황이다. 따라서 생성망은 실제 데이터의 분포를 모두 학습하지 못하여 R2R controller로 사용하기 부적합하다. 도9의 (b)는 hyper-parameters를 BO로 최적화한 후의 생성망과 판별망의 손실 그래프로 손실이 어느 한쪽으로 치우치지 않고 생성망과 판별망이 안정되게 학습되었음을 보여준다. 이러한 안정적인 손실 분포는 R2R controller의 재현성을 확보해주고 새로운 데이터의 빠른 학습을 가능하게 한다.
- [0091] 도10는 hyper-parameter 최적화를 위한 실험 결과를 나타낸 도면이다. 1,450개의 실험 데이터에 대한 hyper-parameter 최적화 알고리즘을 적용한 후의 학습부의 성능 결과를 보여준다. 도10에서 "Random"과 "BO"는 각각 hyper-parameters에 도8의 random values와 BO results를 각각 사용했을 때의 성능 측정값이다. 8개의 control parameter의 결과를 보면 hyper-parameter 최적화 후 성능이 큰 폭으로 개선되는 것을 확인할 수 있었다. MAE측면에서 각 파라미터별 적게는 약 26.4%, 많게는 약 63.8% 정도 개선된 성능을 보였고, +SD의 측면에서도 약

31.2%에서 80.3%까지의 성능 개선이 있었다. 도8과 도9의 결과로부터, B0가 찾은 hyper-parameter solution은 학습부의 생성망과 판별망의 능력에 적절한 균형을 이루는 안정적인 학습을 가능하게 함을 알 수 있다. 또한 B0 결과를 적용한 생성망은 낮은 제어 오차와 변동성을 가진 R2R controller로 사용 가능함을 확인할 수 있다.

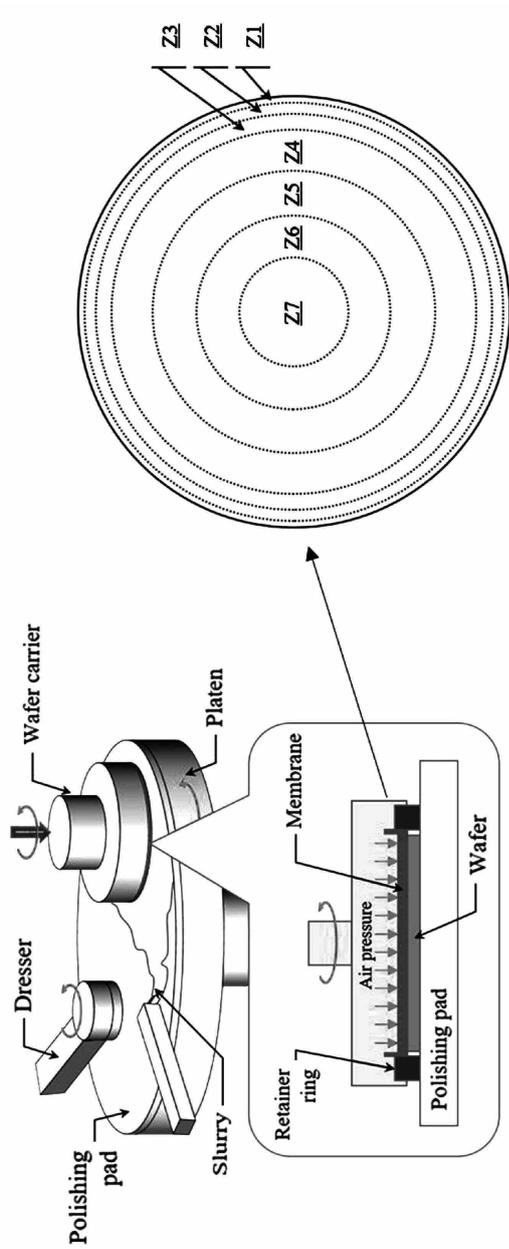
[0092] 도11은 실험용 웨이퍼 공정의 제어 오차를 공정 시퀀스 플롯(run sequence plot)으로 표현한 결과다. 이 플롯에는 CMP 장비 부품의 교체에 의해 발생하는 프로세스 시프트(process shift)가 1회 포함된다. 도11의 (a)와 (b)의 왼쪽 y축은 각각 폴리싱 시간들 ( $X_{T1}, X_{T2}, X_{T3}$ )의 절대 제어 오차의 합과 압력값들 ( $X_{Z1}, X_{Z3}, X_{Z5}, X_{Z6}, X_{Z7}$ )의 절대 제어 오차의 합이며, 오른쪽 y축은 부품들의 정규화된 수명을 나타내며 1의 값에 도달하면 교체에 의해 0에서부터 증가함을 확인할 수 있다.

[0093] 이상에서 설명한 본 발명의 실시예를 구성하는 모든 구성요소들이 하나로 결합하거나 결합하여 동작하는 것으로 기재되어 있다고 해서, 본 발명이 반드시 이러한 실시예에 한정되는 것은 아니다. 즉, 본 발명의 목적 범위 안에서라면, 그 모든 구성요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다. 또한, 그 모든 구성요소들이 각각 하나의 독립적인 하드웨어로 구현될 수 있지만, 각 구성요소들의 그 일부 또는 전부가 선택적으로 조합되어 하나 또는 복수개의 하드웨어에서 조합된 일부 또는 전부의 기능을 수행하는 프로그램 모듈을 갖는 컴퓨터 프로그램으로서 구현될 수도 있다. 또한, 이와 같은 컴퓨터 프로그램은 USB 메모리, CD 디스크, 플래쉬 메모리 등과 같은 컴퓨터가 읽을 수 있는 기록매체(Computer Readable Media)에 저장되어 컴퓨터에 의하여 읽혀지고 실행됨으로써, 본 발명의 실시예를 구현할 수 있다. 컴퓨터 프로그램의 기록매체로서는 자기 기록매체, 광 기록매체 등이 포함될 수 있다.

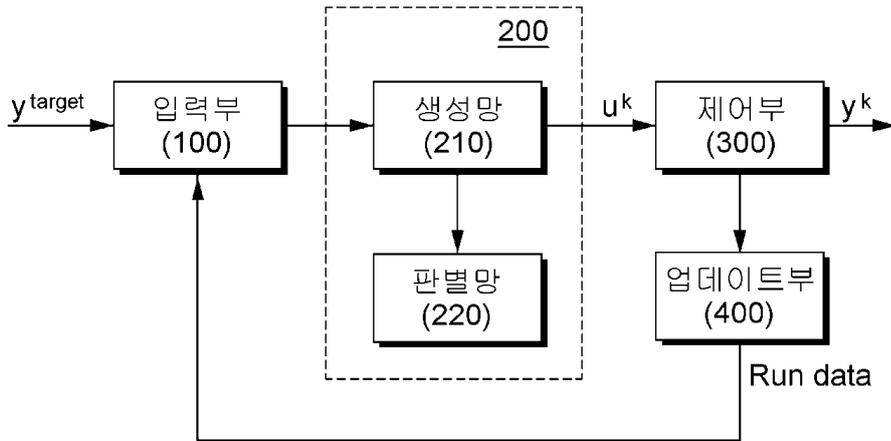
[0094] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경 및 치환이 가능할 것이다. 따라서, 본 발명에 개시된 실시예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구 범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리 범위에 포함되는 것으로 해석되어야 할 것이다.

도면

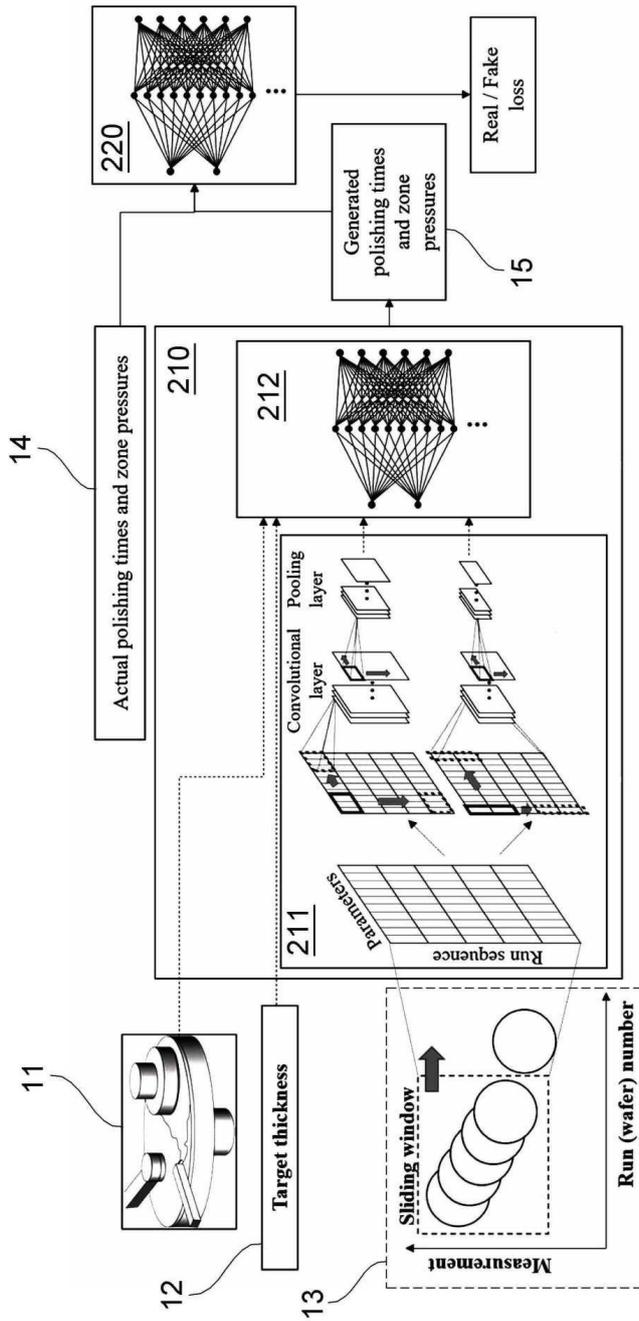
도면1



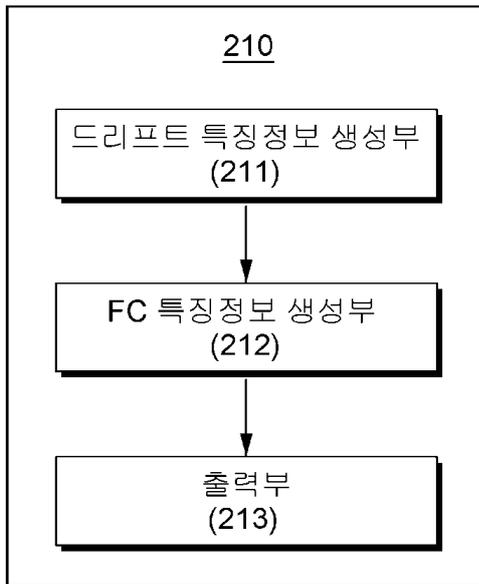
도면2



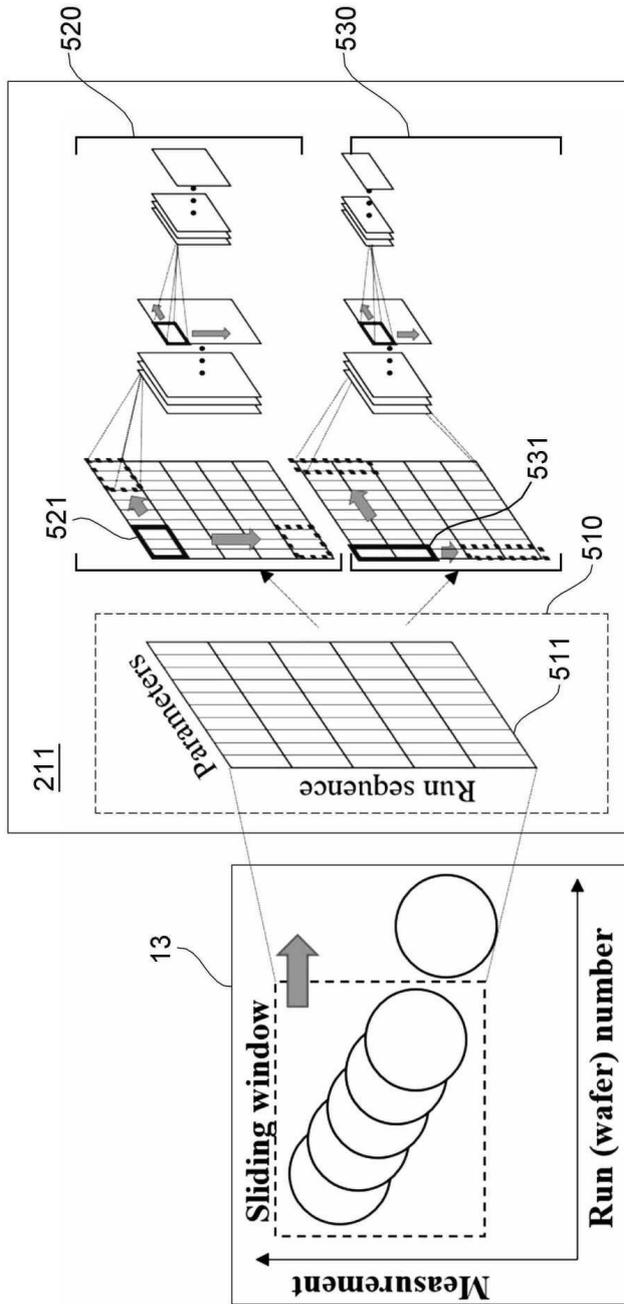
도면3



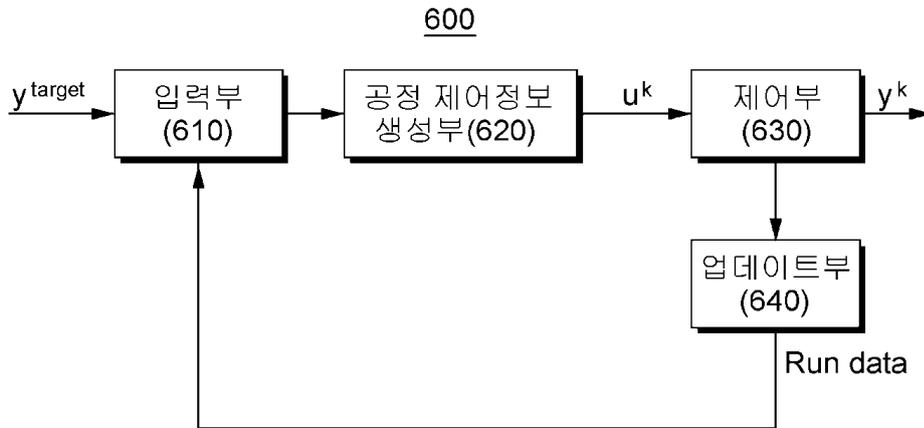
도면4



도면5



도면6



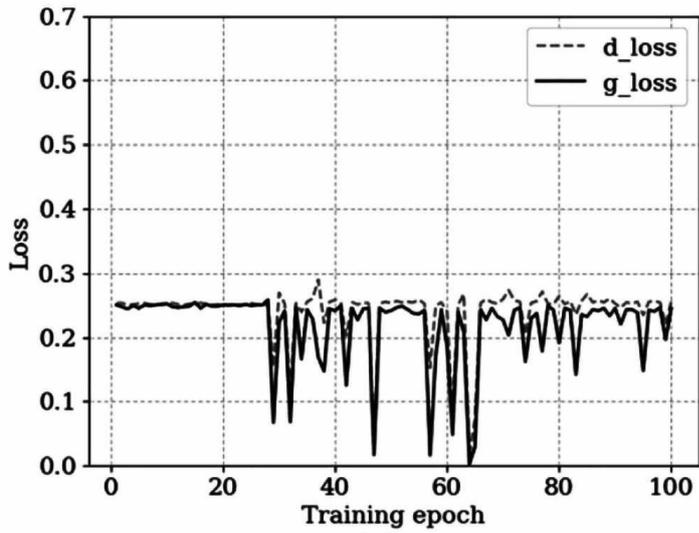
도면7

성능 지표	수식
$ mean_p  + SD_p$ (standard deviation)	$ mean_p  = \left  \frac{1}{n} \sum_{k=1}^n (u_{k,p} - \hat{u}_{k,p}) \right $ $SD_p = \sqrt{\frac{1}{n} \sum_{k=1}^n (u_{k,p} - \bar{u}_p)^2}, \text{ where } \bar{u}_p = \frac{1}{n} \sum_{k=1}^n u_{k,p}$
$MAE_p$ (mean absolute error)	$MAE_p = \frac{1}{n} \sum_{k=1}^n  u_{k,p} - \hat{u}_{k,p} $

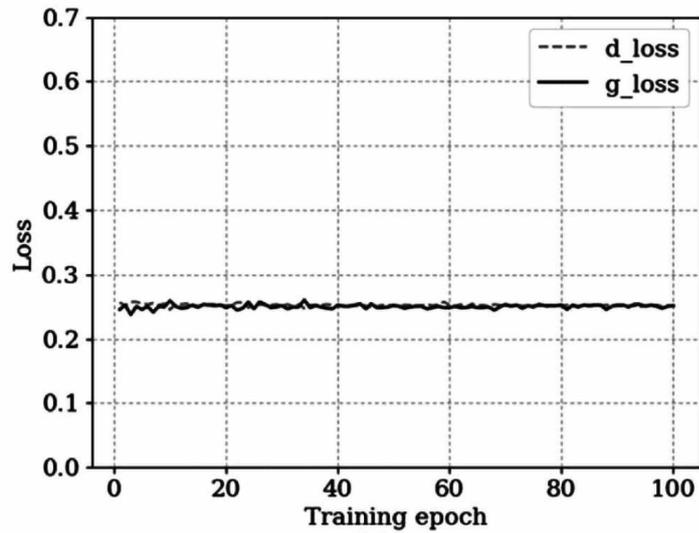
## 도면8

Model	Component	Hyper-parameters	Random search	BO algorithm
(A)		Learning rate	0.005520	<b>0.009962</b>
	VM	Number of feature maps in the 1st CNN	129	<b>92</b>
		Number of feature maps in the 2nd CNN	103	<b>20</b>
		Types of pooling layers (average/max)	Average	<b>Max</b>
		Number of hidden nodes in the 1st FC	361	<b>355</b>
		Number of hidden nodes in the 2nd FC	369	<b>394</b>
		Learning rate	0.000170	<b>0.000001</b>
	R2R controller	Number of hidden nodes in the 1st FC	36	<b>348</b>
		Number of hidden nodes in the 2nd FC	283	<b>366</b>
	(B)		Learning rate	0.002975
Generator		Number of feature maps in the 1st CNN	104	<b>274</b>
		Number of feature maps in the 2nd CNN	192	<b>198</b>
		Types of pooling layers (average/max)	Max	<b>Average</b>
		Number of hidden nodes in the 1st FC	56	<b>242</b>
		Number of hidden nodes in the 2nd FC	127	<b>126</b>
		Learning rate	0.000007	<b>0.000008</b>
Discriminator		Number of hidden nodes in the 1st FC	97	<b>227</b>
		Number of hidden nodes in the 2nd FC	321	<b>302</b>

도면9



(a)



(b)

## 도면10

Measure	$x_{T1}$		$x_{T2}$		$x_{T3}$		$x_{Z1}$	
	Random	BO	Random	BO	Random	BO	Random	BO
mean	-0.52	0.04	0.02	-0.05	-0.35	-0.28	0.05	0.02
SD	2.69	1.49	1.57	1.11	2.64	2.49	0.51	0.36
MAE	2.02	<b>1.18</b>	1.36	<b>0.73</b>	1.76	<b>1.63</b>	0.41	<b>0.27</b>
mean +SD	3.21	<b>1.53</b>	1.59	<b>1.16</b>	2.99	<b>2.76</b>	0.56	<b>0.37</b>
Measure	$x_{Z3}$		$x_{Z5}$		$x_{Z6}$		$x_{Z7}$	
	Random	BO	Random	BO	Random	BO	Random	BO
mean	0.11	0.02	-0.05	-0.01	-0.03	-0.02	-0.12	0.01
SD	0.32	0.33	0.30	0.21	0.40	0.25	0.90	0.49
MAE	0.25	<b>0.25</b>	0.24	<b>0.16</b>	0.32	<b>0.19</b>	0.62	<b>0.31</b>
mean +SD	0.43	<b>0.35</b>	0.35	<b>0.23</b>	0.42	<b>0.27</b>	1.02	<b>0.50</b>

도면11

