(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0135246
(43) 공개일자 2022년10월06일

(51) 국제특허분류(Int. Cl.)
G06N 20/00 (2019.01)
(52) CPC특허분류
G06N 20/00 (2021.08)
(21) 출원번호 10-2022-7030371
(22) 출원일자(국제) 2020년03월31일
심사청구일자 2022년09월01일
(85) 번역문제출일자 2022년09월01일
(86) 국제출원번호 PCT/JP2020/014727
(87) 국제공개번호 WO 2021/199227
국제공개일자 2021년10월07일

(71) 출원인
주식회사 히타치하이테크
일본국 도쿄토 미나토쿠 토라노몬 1쵸메 17방 1고
(72) 발명자
요시다 야스히로
일본 1008280 도꾜 지요다꾸 마루노우찌 1쪼메 6-6 가부시키가이샤 히타치 세이사쿠쇼 내
이시까와 마사요시
일본 1008280 도꾜 지요다꾸 마루노우찌 1쪼메 6-6 가부시키가이샤 히타치 세이사쿠쇼 내
(뒷면에 계속)
(74) 대리인
장수길, 박상돈, 이중희

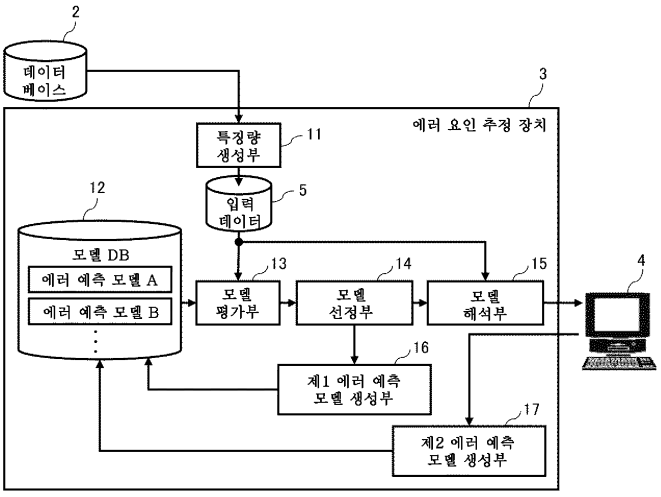
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 **에러 요인의 추정 장치 및 추정 방법**

(57) 요 약

에러 요인의 추정 장치는, 외부로부터 보내진 데이터를 사용하여 기계 학습 모델에 적합한 특징량을 생성하는 특징량 생성부와, 특징량을 입력 데이터로서 에러 발생의 유무의 판정에 사용되는 에러 예측 모델을 적어도 하나 이상 갖는 모델 데이터베이스와, 에러 예측 모델의 예측 결과와 실제로 계측된 참된 에러 결과를 비교하여 에러 예측 모델의 성능을 평가하는 모델 평가부와, 모델 평가부에서 계산된 평가값이 미리 설정된 규정값 이상이 되는 에러 예측 모델을 모델 데이터베이스로부터 선정하는 모델 선정부와, 모델 선정부에서 선택되는 에러 예측 모델의 해당이 없는 경우에는, 계측된 에러에 대하여 새로운 에러 예측 모델을 생성하는 에러 예측 모델 생성부를 구비하고 있다. 이에 의해, 에러 요인의 사전의 애노테이션이 없어도, 발생하는 다종류의 에러에 대하여 그 요인을 추정할 수 있다.

대 표 도 - 도2



(72) 발명자

하야까와 고우이찌

일본 1056409 도꾜 미나또꾸 도라노몬 1쪼메 17-1 주식회사 히타치 하이테크 내

다까노 마사미

일본 1056409 도꾜 미나또꾸 도라노몬 1쪼메 17-1 주식회사 히타치 하이테크 내

사사지마 후미히로

일본 1056409 도꾜 미나또꾸 도라노몬 1쪼메 17-1 주식회사 히타치 하이테크 내

명 세 서

청구범위

청구항 1

외부로부터 보내진 데이터를 사용하여 기계 학습 모델에 적합한 특징량을 생성하는 특징량 생성부와,

상기 특징량을 입력 데이터로서 에러 발생의 유무의 판정에 사용되는 에러 예측 모델을 적어도 하나 이상 갖는 모델 데이터베이스와,

상기 에러 예측 모델의 예측 결과와 실제로 계측된 참된 에러 결과를 비교하여 상기 에러 예측 모델의 성능을 평가하는 모델 평가부와,

상기 모델 평가부에서 계산된 평가값이 미리 설정된 규정값 이상이 되는 상기 에러 예측 모델을 상기 모델 데이터베이스로부터 선정하는 모델 선정부와,

상기 모델 선정부에서 선택되는 상기 에러 예측 모델의 해당이 없는 경우에는, 계측된 에러에 대하여 새로운 에러 예측 모델을 생성하는 에러 예측 모델 생성부를 구비한, 에러 요인의 추정 장치.

청구항 2

제1항에 있어서, 상기 입력 데이터 내의 에러 데이터를 에러 요인마다 분류하는 데이터 분류부를 더 구비한, 에러 요인의 추정 장치.

청구항 3

제2항에 있어서, 상기 에러 예측 모델 생성부는, 분류된 상기 에러 데이터마다 다른 라벨 부여를 하여, 이 라벨과 함께 상기 에러 예측 모델을 생성하고, 이 에러 예측 모델을 상기 모델 데이터베이스에 송신하는, 에러 요인의 추정 장치.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 모델 선정부에서 선정된 상기 에러 예측 모델에 있어서의 에러 판정 결과에 대한 상기 특징량의 기여도를 수치화하는 모델 해석부를 더 구비한, 에러 요인의 추정 장치.

청구항 5

제4항에 있어서, 상기 모델 해석부에서 계산된 상기 기여도의 값이 높은 것의 상기 특징량에 대해서, 에러 요인 후보로서 유저에게 제시하는 구성을 갖는, 에러 요인의 추정 장치.

청구항 6

제4항에 있어서, 상기 모델 선정부에서 복수의 에러 예측 모델이 선정된 경우에는,

상기 모델 평가부는, 모델 평가값을 산출하고,

상기 모델 해석부에서 계산된 상기 특징량마다의 상기 기여도는, 상기 모델 평가값을 사용하여 보정되고, 상기 복수의 에러 예측 모델의 각각으로부터 계산된 보정 후의 기여도의 값이 높은 것의 상기 특징량에 대해서, 에러 요인 후보로서 유저에게 제시하는 구성을 갖는, 에러 요인의 추정 장치.

청구항 7

제5항 또는 제6항에 있어서, 상기 에러 요인 후보에 대하여 상기 유저로부터 수정이 있었던 경우에는, 수정 후의 에러 요인이 상기 모델 해석부의 해석 결과에 포함되도록 에러 예측 모델을 생성하는 또 하나의 에러 예측 모델 생성부를 더 구비한, 에러 요인의 추정 장치.

청구항 8

제4항에 있어서, 상기 특징량 생성부에서 생성된 상기 특징량 및 그 조합 중 적어도 어느 한쪽에 대응하는 에러

요인의 관계를 보존하는 에러 요인 라벨 데이터베이스와,

상기 모델 해석부에서 수치화된 상기 기여도에 대응하는 상기 특징량에 대하여 상기 에러 요인 라벨 데이터베이스 내의 라벨 관계를 사용하여 대응하는 에러 요인을 부여하는 에러 요인 라벨 취득부를 더 구비한, 에러 요인의 추정 장치.

청구항 9

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 에러 예측 모델 생성부는, 대상으로 하는 에러가 발생한 동작 공정 및 그 이전의 동작 공정에서의 상기 입력 데이터를 사용하여 새로운 에러 예측 모델을 생성하는, 에러 요인의 추정 장치.

청구항 10

외부로부터 보내진 데이터를 사용하여 기계 학습 모델에 적합한 특징량을 생성하는 특징량 생성 공정과,

모델 데이터베이스에 보존되어 있는 에러 예측 모델의 예측 결과와 실제로 계측된 참된 에러 결과를 비교하여, 상기 특징량을 입력 데이터로서 에러 발생의 유무의 판정에 사용되는 에러 예측 모델의 성능을 평가하는 모델 평가 공정과,

상기 모델 평가 공정에서 계산된 평가값이 미리 설정된 규정값 이상이 되는 상기 에러 예측 모델을 상기 모델 데이터베이스로부터 선정하는 모델 선정 공정과,

상기 모델 선정 공정에서 선택되는 상기 에러 예측 모델의 해당이 없는 경우에는, 계측된 에러에 대하여 새로운 에러 예측 모델을 생성하는 에러 예측 모델 생성 공정을 포함하는, 에러 요인의 추정 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은, 에러 요인의 추정 장치 및 추정 방법에 관한 것이다.

배 경 기 술

[0002] 반도체 계측 장치나 반도체 검사 장치는, 레시피라고 불리는 설정 파라미터에 따라, 반도체 웨이퍼의 표면에 있어서 이상이라고 판정된 검사점마다 검사 동작이나 계측 동작을 실시한다. 레시피 파라미터의 조정은, 계측·검사 대상의 속성이나 장치의 특성 등에 따라, 엔지니어가 매뉴얼 작업에 의해 각 항목을 최적화하는 것이 일반적이다. 따라서, 예를 들어 조정이 충분하지 않은 레시피의 사용이나, 경시 변화에 의해 장치의 특성이 바뀐 경우 등에 있어서는, 검사 동작이나 계측 동작에 있어서 에러가 발생할 가능성이 있다. 이러한 에러는, 레시피의 내용에 기인하는 에러로서 레시피 에러라고 불린다.

[0003] 이러한 레시피 에러가 발생했을 때에는, 서비스 엔지니어가 반도체 계측 장치나 반도체 검사 장치로부터 장치 내부 데이터를 해석하여 원인 개소를 특정하는 것이 일반적이다. 그러나, 반도체의 미세화·다양화에 수반하여, 레시피 수 및 레시피 설정 항목의 증가, 레시피 작성의 복잡화 등이 발생하고 있다. 이 때문에, 레시피 에러의 원인 특정에는 시간을 요하고, 장치의 가동률을 저하시키는 하나의 요인이 되고 있다.

[0004] 특허문헌 1에는, 초소형 전자 기구의 원하는 치수를 측정하는데 사용되는 계측 툴에 있어서의 장해를 식별하는 방법에 의해, 유저가, 가장 문제가 있는 레시피에 신속히 집중하고, 어느 계측 툴에도 전형적으로 존재하는 에러 로그를 사용하여 근본적인 원인을 결정할 수 있는 것, 이 처리를 자동화할 수 있는 것 등이 개시되어 있다.

[0005] 특허문헌 2에는, 워크의 가공면에 불량이 발생한 경우에 그 요인을 추정하는 기술로서, 기계 학습 장치를 사용하여, 검사 장치에 의한 워크의 가공면의 검사 결과를 상태 변수로서 관측하고, 가공면 불량의 발생 요인을 나타내는 라벨 데이터를 취득하여, 상태 변수와 라벨 데이터를 관련지어서 학습하는, 가공 불량 요인 추정 장치가 개시되어 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 일본 특허 제4398441호 공보

(특허문헌 0002) 일본 특허 제6530779호 공보

발명의 내용

해결하려는 과제

[0007] 특허문헌 1에 기재된 방법에 있어서는, 전형적인 에러 로그를 사용하여 근본적인 원인을 자동적으로 결정할 수 있다고 되어 있다. 그러나, 특허문헌 1에 있어서는, 계측 툴에 의해 사용된 레시피에 관한 정규화된 수의 에러가 구체적으로 어떤 에러인지 등이 기재되어 있지 않다.

[0008] 특허문헌 2에 기재된 가공 불량 요인 추정 장치의 적용 범위는, 상태 변수와 라벨 데이터를 관련지어서 학습할 수 있는 경우에 한정된다. 바꾸어 말하면, 애노테이션이 필요해진다.

[0009] 반도체 검사 장치 등에 있어서의 레시피 에러의 요인 추정 시에는, 에러가 어떤 메커니즘으로 발생하는지는, 제품·제조 프로세스 의존이다. 또한, 에러의 베리에이션은 다양하다. 이 때문에, 에러 요인을 사전에 상정하고 망라하여 학습에 사용하는 것은 곤란하다.

[0010] 본 발명은, 에러 요인의 사전의 애노테이션이 없어도, 발생하는 다종류의 에러에 대하여 그 요인을 추정하는 것을 목적으로 한다.

과제의 해결 수단

[0011] 본 발명의 에러 요인의 추정 장치는, 외부로부터 보내진 데이터를 사용하여 기계 학습 모델에 적합한 특징량을 생성하는 특징량 생성부와, 특징량을 입력 데이터로서 에러 발생의 유무의 판정에 사용되는 에러 예측 모델을 적어도 하나 이상 갖는 모델 데이터베이스와, 에러 예측 모델의 예측 결과와 실제로 계측된 참된 에러 결과를 비교하여 에러 예측 모델의 성능을 평가하는 모델 평가부와, 모델 평가부에서 계산된 평가값이 미리 설정된 규정값 이상이 되는 에러 예측 모델을 모델 데이터베이스로부터 선정하는 모델 선정부와, 모델 선정부에서 선택되는 에러 예측 모델의 해당이 없는 경우에는, 계측된 에러에 대하여 새로운 에러 예측 모델을 생성하는 에러 예측 모델 생성부를 구비하고 있다.

[0012] 본 발명의 에러 요인의 추정 방법은, 외부로부터 보내진 데이터를 사용하여 기계 학습 모델에 적합한 특징량을 생성하는 특징량 생성 공정과, 모델 데이터베이스에 보존되어 있는 에러 예측 모델의 예측 결과와 실제로 계측된 참된 에러 결과를 비교하여, 특징량을 입력 데이터로서 에러 발생의 유무의 판정에 사용되는 에러 예측 모델의 성능을 평가하는 모델 평가 공정과, 모델 평가 공정에서 계산된 평가값이 미리 설정된 규정값 이상이 되는 에러 예측 모델을 모델 데이터베이스로부터 선정하는 모델 선정 공정과, 모델 선정 공정에서 선택되는 에러 예측 모델의 해당이 없는 경우에는, 계측된 에러에 대하여 새로운 에러 예측 모델을 생성하는 에러 예측 모델 생성 공정을 포함한다.

발명의 효과

[0013] 본 발명에 따르면, 에러 요인의 사전의 애노테이션이 없어도, 발생하는 다종류의 에러에 대하여 그 요인을 추정할 수 있다.

도면의 간단한 설명

[0014] 도 1은, 실시예 1에 관한 에러 요인 추정 장치를 포함하는 정보 처리 시스템을 도시하는 블록도이다.

도 2는, 도 1의 에러 요인 추정 장치를 도시하는 구성도이다.

도 3은, 실시예 1에 관한 제1 에러 예측 모델 생성부에 있어서의 새로운 에러 예측 모델의 생성 수순을 도시하는 흐름도이다.

도 4는, 실시예 1에 관한 입력 데이터 데이터 구조의 일례를 도시하는 표이다.

도 5는, 도 2의 모델 해석부에 있어서의 처리 및 표시예를 도시하는 모식도이다.

도 6은, 실시예 1에 관한 모델 평가부 및 모델 해석부에 있어서의 처리 및 표시예를 도시하는 모식도이다.

도 7은, 실시예 1에 관한 특징량의 값과 그 에러 발생에 대한 기여도의 관계의 일례를 도시하는 그래프이다.

도 8은, 실시예 1에 관한 제2 에러 예측 모델 생성 수순을 도시하는 흐름도이다.

도 9는, 실시예 2에 관한 데이터 분류부를 구비한 에러 요인 추정 장치를 도시하는 구성도이다.

도 10은, 실시예 2에 관한 데이터 분류부를 도시하는 구성도이다.

도 11은, 특징량의 값과 그 에러 발생에 대한 기여도의 관계를 사용하여 에러 데이터를 분류한 상태를 도시하는 그래프이다.

도 12는, 실시예 3에 관한 에러 요인 추정만을 실시하는 에러 요인 추정 장치를 도시하는 구성도이다.

도 13은, 실시예 3에 관한 에러 요인 라벨 취득부 및 에러 요인 라벨 데이터베이스를 구비한 에러 요인 추정 장치를 도시하는 구성도이다.

도 14는, 실시예 3에 관한 유저에 대한 에러 요인 후보를 도시하는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0015] 이하에 설명하는 실시 형태에 있어서, 「반도체 검사 장치」는, 반도체 웨이퍼의 표면에 형성된 패턴의 치수를 계측하는 장치, 반도체 웨이퍼의 표면에 형성된 패턴의 결함 유무를 검사하는 장치, 혹은 패턴이 형성되어 있지 않은 베어 웨이퍼의 결함의 유무를 검사하는 장치 등을 포함하고, 이들의 장치를 복수 조합한 복합 장치도 포함하는 것이다.

[0016] 또한, 이하에 설명하는 실시 형태에 있어서, 「검사」란, 계측 또는 검사의 의미로 사용하는 것으로 하고, 「검사 동작」이란, 계측 동작 또는 검사 동작의 의미로 사용하는 것으로 하고, 「검사 대상」이란, 계측 또는 검사의 대상으로 되는 웨이퍼, 혹은, 당해 웨이퍼에 있어서의 계측 또는 검사의 대상 영역을 가리키는 것으로 한다.

[0017] 또한, 본 명세서에 있어서, 에러 요인 추정 장치는, 에러 요인의 추정 장치와 동의이고, 에러 요인 추정 방법은, 에러 요인의 추정 방법과 동의이다.

[0018] 이하, 에러 요인의 추정 장치 및 추정 방법의 바람직한 실시 형태에 대하여 설명한다.

[0019] 상기 추정 장치는, 특징량 생성부와, 모델 데이터베이스와, 모델 평가부와, 모델 선정부와, 에러 예측 모델 생성부를 구비하고, 입력 데이터 내의 에러 데이터를 에러 요인마다 분류하는 데이터 분류부를 더 구비하고 있는 것이 바람직하다.

[0020] 상기 추정 장치에 있어서, 에러 예측 모델 생성부는, 분류된 에러 데이터마다 다른 라벨 부여를 하여, 이 라벨과 함께 에러 예측 모델을 생성하고, 이 에러 예측 모델을 모델 데이터베이스에 송신하는 것이 바람직하다.

[0021] 상기 추정 장치는, 모델 선정부에서 선정된 에러 예측 모델에 있어서의 에러 판정 결과에 대한 특징량의 기여도를 수치화하는 모델 해석부를 더 구비하고 있는 것이 바람직하다.

[0022] 상기 추정 장치는, 모델 해석부에서 계산된 기여도의 값이 높은 것의 특징량에 대해서, 에러 요인 후보로서 유저에게 제시하는 구성을 갖는 것이 바람직하다.

[0023] 상기 추정 장치에 있어서, 모델 선정부에서 복수의 에러 예측 모델이 선정된 경우에는, 모델 평가부는, 모델 평가값을 산출하고, 모델 해석부에서 계산된 특징량마다의 기여도는, 모델 평가값을 사용하여 보정되고, 복수의 에러 예측 모델의 각각으로부터 계산된 보정 후의 기여도의 값이 높은 것의 특징량에 대해서, 에러 요인 후보로서 유저에게 제시하는 구성을 갖는 것이 바람직하다.

[0024] 상기 추정 장치는, 에러 요인 후보에 대하여 유저로부터 수정이 있었던 경우에는, 수정 후의 에러 요인이 모델 해석부의 해석 결과에 포함되도록 에러 예측 모델을 생성하는 또 하나의 에러 예측 모델 생성부를 더 구비하고 있는 것이 바람직하다.

[0025] 상기 추정 장치는, 특징량 생성부에서 생성된 특징량 및 그 조합 중 적어도 어느 한쪽에 대응하는 에러 요인의 관계를 보존하는 에러 요인 라벨 데이터베이스와, 모델 해석부에서 수치화된 기여도에 대응하는 특징량에 대하여 에러 요인 라벨 데이터베이스 내의 라벨 관계를 사용하여 대응하는 에러 요인을 부여하는 에러 요인 라벨 취득부를 더 구비하고 있는 것이 바람직하다.

[0026] 상기 추정 장치에 있어서, 에러 예측 모델 생성부는, 대상으로 하는 에러가 발생한 동작 공정 및 그 이전의 동작 공정에서의 입력 데이터를 사용하여 새로운 에러 예측 모델을 생성하는 것이 바람직하다.

[0027] 또한, 추정 장치의 구성과 추정 방법의 공정의 관계에 대해서, 특징량 생성부는 특징량 생성 공정에, 모델 평가부는 모델 평가 공정에, 모델 선정부는 모델 선정 공정에, 에러 예측 모델 생성부는 에러 예측 모델 생성 공정에 대응한다. 또한, 이들의 공정은, 하나의 장치에 있어서 실시되는 것에 한정되는 것은 아니고, 분산 배치된 복수의 장치에 의해 실시되어도 된다.

[0028] **실시예 1**

[0029] 도 1은, 실시예 1에 관한 에러 요인 추정 장치를 포함하는 정보 처리 시스템의 일례를 도시한 것이다.

[0030] 본 도면에 있어서, 반도체 검사 장치(1)는, 네트워크(101)를 통해 데이터베이스(2) 및 에러 요인 추정 장치(3)에 접속되어 있다. 에러 요인 추정 장치(3)는, 단말기(4)(GUI)에 접속되어 있다. 에러 요인 추정 장치(3)는, 반도체 검사 장치(1)가 실시하는 검사 동작에 있어서의 에러의 요인을 추정한다.

[0031] 반도체 검사 장치(1)로부터 보내지는 데이터에는, 예를 들어 장치 데이터, 계측 레시피(이하 단순히 「레시피」라고 칭하는 경우도 있음), 계측 결과 및 에러 결과가 포함된다. 또한, 레시피에는, 계측점 수, 계측점(Evaluation Point: EP)의 좌표 정보, 화상을 촬상할 때의 촬상 조건, 촬상 시퀀스 등이 포함되어 있어도 된다. 또한, 레시피에는, 계측점에 맞춰서 계측점을 계측하기 위한 준비 단계에서 취득되는 화상의 좌표나 촬상 조건 등이 포함되어 있어도 된다.

[0032] 장치 데이터는, 장치 고유 파라미터, 장치 기차 보정 데이터 및 관찰 조건 파라미터를 포함한다. 장치 고유 파라미터는, 반도체 검사 장치(1)를 규정 사양대로 동작시키기 위하여 사용하는 보정 파라미터이다. 장치 기차 보정 데이터는, 반도체 검사 장치 간의 기차를 보정하기 위하여 사용하는 파라미터이다. 관찰 조건 파라미터는, 예를 들어 전자 광학계의 가속 전압 등의 주사형 전자 현미경(SEM)의 관찰 조건을 규정하는 파라미터이다.

[0033] 레시피는, 레시피 파라미터로서, 웨이퍼 맵, 얼라인먼트 파라미터, 어드레싱 파라미터 및 측장 파라미터를 포함한다. 웨이퍼 맵은, 반도체 웨이퍼의 표면의 좌표 맵(예를 들어 패턴의 좌표)이다. 얼라인먼트 파라미터는, 예를 들어 반도체 웨이퍼의 표면의 좌표계와 반도체 검사 장치(1) 내부의 좌표계와의 사이의 어긋남을 보정하기 위하여 사용하는 파라미터이다. 어드레싱 파라미터는, 예를 들어 반도체 웨이퍼의 표면에 형성되어 있는 패턴 중, 검사 대상 영역 내에 존재하는 특징적인 패턴을 특정하는 정보이다. 측장 파라미터는, 길이를 측정하는 조건을 기술한 파라미터이고, 예를 들어 패턴 중 어느 부위의 길이를 측정할지를 지정하는 파라미터이다.

[0034] 계측 결과는, 측장 결과, 화상 데이터 및 동작 로그를 포함한다. 측장 결과는, 반도체 웨이퍼의 표면의 패턴 길이를 측정한 결과를 기술한 것이다. 화상 데이터는, 반도체 웨이퍼의 관찰 화상이다. 동작 로그는, 얼라인먼트, 어드레싱 및 측장의 각 동작 공정에서의 반도체 검사 장치(1)의 내부 상태를 기술한 데이터이다. 예를 들어, 각 부품의 동작 전압, 관찰 시야의 좌표 등을 들 수 있다.

[0035] 에러 결과는, 에러가 발생한 경우에, 얼라인먼트, 어드레싱 및 측장의 각 동작 공정의 어느 것에서 발생한 에러인지를 나타내는 파라미터이다.

[0036] 이들의 장치 데이터, 레시피, 계측 결과, 에러 결과 등의 데이터는, 네트워크(101)를 통해 데이터베이스(2)에 축적된다. 축적된 데이터는, 에러 요인 추정 장치(3)에서 해석된다. 해석 결과는, 단말기(4)에서 유저가 판독할 수 있는 형식으로 표시된다.

[0037] 도 2는, 도 1의 에러 요인 추정 장치의 상세한 구성을 도시한 것이다.

[0038] 도 2에 있어서, 에러 요인 추정 장치(3)는, 외부의 데이터베이스(2)에 접속된 특징량 생성부(11)와, 입력 데이터 기록부(5)와, 모델 데이터베이스(12)(모델 DB)와, 모델 평가부(13)와, 모델 선정부(14)와, 모델 해석부(15)와, 제1 에러 예측 모델 생성부(16)와, 제2 에러 예측 모델 생성부(17)를 갖고 있다. 또한, 제1 에러 예측 모델 생성부(16)는, 단순히 「에러 예측 모델 생성부」라고도 칭한다. 또한, 제2 에러 예측 모델 생성부(17)는, 「또 하나의 에러 예측 모델 생성부」라고도 칭한다.

[0039] 특징량 생성부(11)는, 데이터베이스(2)로부터 보내진 장치 데이터, 레시피, 계측 결과 등의 생 데이터로부터, 기계 학습 모델에 적합한 특징량을 추출하고, 당해 특징량을 입력 데이터 기록부(5)에 출력한다. 여기서, 특징량의 추출은, 데이터의 스케일링, 카테고리컬 변수의 인코딩, 교호 작용 특징량 등의 복수의 데이터를 조합한

복합적인 특징량 작성 등을 포함하는 것이어도 된다.

[0040] 모델 데이터베이스(12)에는, 입력 데이터 기록부(5) 내의 데이터를 입력으로 하여, 각 검사점에 있어서의 에러 발생의 유무의 판정에 사용되는 에러 예측 모델을 미리 적어도 하나 이상 기록해 둔다. 이 미리 기록된 초기의 에러 예측 모델에 대해서는, 다른 반도체 제조 공장이나 제조 라인에서 생성된 모델을 전용해도 되고, 데이터베이스(2) 내의 임의의 에러에 대하여 후술하는 모델 생성 수순에 기초하여 구축해도 된다.

[0041] 모델 평가부(13)는, 입력 데이터 기록부(5) 내의, 예를 들어 레시피 단위, 웨이퍼 단위, 검사점 단위 등의 데이터에 대하여 모델 데이터베이스(12) 내의 에러 예측 모델의 성능을 평가한다. 성능 평가는, 에러 예측 모델을 사용하여 판정된 에러 예측 결과와, 입력 데이터 기록부(5) 내의 참된 에러 결과를 비교함으로써 얻어진다. 성능의 평가값으로서는, 정밀도, 재현율, 적합율, F1값, AUC 등을 사용할 수 있다. 여기서, F1값은, 적합율과 재현율의 조화 평균이다. 또한, AUC는, Area Under the Curve의 약칭이다.

[0042] 모델 선정부(14)는, 모델 평가부(13)에 있어서 평가값이 높은 모델을, 입력 데이터 기록부(5)에 포함되는 에러의 판정에 적합한 모델로서 하나 이상 선정한다. 모델의 선정 방법으로서는, 모델 평가부(13)에서 사용한 평가값에 대하여 미리 규정값을 설정하고, 이 규정값 이상이 되는 평가값의 모델 중에서 선정한다.

[0043] 모델 선정부(14)에 있어서, 설정된 규정값 이상의 평가값이 된 모델이 존재하지 않는 경우에는, 생성 완료된 에러 예측 모델에 적합하지 않은 신규 에러가 입력된 것으로 해서, 제1 에러 예측 모델 생성부(16)에서 새로운 에러 예측 모델을 생성한다.

[0044] 모델 해석부(15)는, 모델 선정부(14)에서 선정된 에러 예측 모델에 대해서, 입력 데이터 기록부(5) 내의 각 특징량이 에러 판정에 대하여 어느 정도 기여했는지를 해석함으로써, 에러와 높은 상관을 나타내는 특징량을 추출한다.

[0045] 도 3은, 도 2의 제1 에러 예측 모델 생성부(16)에 있어서의 새로운 에러 예측 모델의 생성 수순을 도시한 것이다.

[0046] 스텝 S100에 있어서, 상기한 바와 같이, 설정된 규정값 이상의 평가값이 된 모델이 존재하지 않는다고 판정된 경우, 스텝 S101로 이행한다.

[0047] 스텝 S101에 있어서는, 에러 예측 모델의 생성에 필요한 학습 데이터의 선정(추출)을 한다. 선정하는 방법으로서는, 에러 예측 모델에서 검출할 수 없었던 에러와 동일한 레시피 또는 유사한 레시피를 포함하는 데이터를 데이터베이스(2) 또는 입력 데이터 기록부(5)로부터 추출한다.

[0048] 이어서, 스텝 S102에서는, 학습 데이터 중 어느 특징량을 우선적으로 사용하는지의 가중을 수정한다. 수정 방법으로서는, 예를 들어 랜덤 서치나 베이즈 최적화 등의 기존의 파라미터 탐색 방법을 활용할 수 있다.

[0049] 스텝 S103에서는, 학습 데이터를 입력으로서 스텝 S102에서 산출한 가중치에 기초하여, 학습 데이터 중에 포함되는 에러 발생의 유무의 판정에 사용되는 학습 모델인 에러 예측 모델을 생성한다. 이 에러 예측 모델은, 결정목이나 Neural Network 등, 어떤 기계 학습 알고리즘을 사용하여 생성한다고 해도 된다.

[0050] 스텝 S104에서는, 스텝 S103에서 생성된 에러 예측 모델의 성능을 평가한다. 평가 방법으로서는, 모델 평가부(13)(도 2)와 마찬가지로, 정밀도, 재현율, 적합률, F1값, AUC 등의 지표를 사용할 수 있다. 또한, 바람직하게는, 교차 검증 등의 방법을 사용하여 이들의 평가값을 계산할 수 있으면 된다.

[0051] 스텝 S105에서는, 스텝 S104에서 계산된 평가값이 미리 정해진 규정값 이상인지의 여부를 판정한다. 규정값 미만이면, 스텝 S102로 되돌아가고, 다시 마찬가지의 처리를 반복하고, 규정값 이상이면 신규 에러 모델의 생성 완료로서 모델 데이터베이스(12)(도 2)에 보존한다.

[0052] 또한, 상기의 유사한 레시피의 선정 방법으로서는, 예를 들어 반도체 웨이퍼의 표면에 형성되어 있는 패턴의 등록 정보를 나타내는 파라미터나 측정 배율의 값이 가까운 것을 선정하면 할 수 있다. 데이터베이스(2)로부터 추출하는 경우에는, 특징량 생성부(11)에서 기기 학습 모델에 적합한 형식으로 특징량을 생성한다. 또한, 추출하는 데이터의 기간을 지정하는 것으로 해도 된다. 과거의 데이터이면, 웨이퍼의 제조 조건이나 장치의 상태가 변화한 가능성이 있기 때문에, 에러 발생 시점으로부터 거슬러 올라서 데이터를 추출하는 기간이 지정되는 것이 바람직하다.

[0053] 또한, 학습 데이터는, 대상(예측 대상)으로 하는 에러가 발생한 동작 공정과, 그 이전의 동작 공정에서의 레시피 또는 계측 결과를 포함하는 것으로 하면 된다. 반도체 검사 장치(1)(도 1)에 있어서의 측정은, 얼라인먼트,

어드레싱, 측장 등이 연속된 동작 공정을 포함한다.

[0054] 도 4는, 반도체 검사 장치에 있어서의 측정으로부터 얻어진 특징량의 값을 포함하는 입력 데이터 데이터 구조의 일례를 도시하는 표이다.

[0055] 본 도면에 있어서는, 각각의 측정 Index에 대한 특징량(Z_1, Z_2, …, Z_m)의 값, 동작 공정 및 에러의 발생 유무(에러 결과)를 나타내고 있다.

[0056] 본 도면에 있어서는, 측정 Index가 2인 경우이며 동작 공정이 2인 경우에 있어서 에러가 발생하고 있다. 이 경우, 그 동작 공정 2보다 뒤의 동작 공정 3 이후의 특징량은, 에러 발생과는 무관계로서 제외해도 되지만, 앞의 동작 공정 1에 대해서는, 에러에 영향을 주고 있을 가능성이 있기 때문에, 학습 데이터에 포함하는 것으로 한다.

[0057] 이와 같이, 학습 완료의 모델에서 대응할 수 없는 에러가 발생한 경우에는, 신규 에러로서 새로운 예측 모델을 생성함으로써, 사전의 애노테이션을 행하지 않고, 에러 요인마다 에러 예측 모델의 생성을 할 수 있다.

[0058] 다음으로, 모델 해석부(15)에 있어서의 에러 판정에 대한 기여도의 계산 방법과, 계산 결과의 유저에 대한 가시화 방법에 대하여 설명한다.

[0059] 도 5는, 에러 예측 모델에 대한 입력된 특징량의 에러 판정에 대한 기여도의 계산과, 유저에 대한 가시화 방법을 도시하는 모식도이다.

[0060] 본 도면에 있어서, 입력 데이터 기록부(5)의 입력 데이터와 모델 데이터베이스(12)(도 2)의 에러 예측 모델 A(210)가 모델 해석부(15)에 입력된다. 그리고, 기여도의 계산 결과(220)가 출력되고, 각각의 특징량의 에러에 대한 기여도의 그래프(230)가 표시 가능하게 된다. 바꾸어 말하면, 특징량의 기여도가 수치화된다. 여기에서는, 입력 데이터 기록부(5)의 입력 데이터는, 특징량(Z_1, Z_2, …, Z_m)의 값 및 에러 발생 결과가, 얼라인먼트, 어드레싱, 측장의 각 동작 공정 때마다 할당되는 측정 Index마다 저장된 데이터 구조를 갖는 예를 나타내고 있다.

[0061] 이들의 특징량의 에러 예측 모델에 있어서의 에러 판정 결과에 대한 기여도는, 예를 들어 에러 예측 모델이 결정목을 베이스로 한 알고리즘에서 구축되는 경우, 각 특징량이 모델 내의 분기에 출현하는 개수나 목적 함수의 개선값 등에 기초하여 계산되는 변수 중요도(Feature Importance)나, 각 특징량의 값의 모델 출력에 대한 감도를 계산하는 SHAP값에 의해 평가할 수 있다. 여기서, SHAP는, 모델의 예측 결과에 대한 각 변수(특징량)의 기여를 구하기 위한 방법이고, SHapley Additive exPlanations의 약칭이다.

[0062] 도 2의 모델 선정부(14)에 있어서 선정된 모델이 에러 예측 모델 A였던 경우, 에러 예측 모델 A의 입력에, 예를 들어 측정 Index가 1인 행의 데이터가 입력되고, 특징량 Z_1을 포함시킨 경우와 제외한 경우의 에러 예측 모델 A의 출력 결과의 차분에 기초하여, 에러 판정에 대한 ϕ 기여도 $\phi_{1,1}$이 계산된다. 이 계산을 모든 특징량 및 모든 측정 Index에 대하여 반복함으로써, 예측 모델의 판정 결과에 대하여 각 특징량이 어느 정도의 영향을 주었는지를 수치화할 수 있다. 각 특징량에 있어서의 SHAP값의 총합 ϕ_{SUM}은, 값이 큰 특징량으로부터 순으로 단말기(4)에 표시할 수 있다. 바꾸어 말하면, 모델 해석부(15)에서 계산된 기여도의 값이 높은 것의 특징량을 에러 요인 후보로 하고, 단말기(4) 등을 통해 유저에게 제시할 수 있다.

[0063] 이와 같이, 입력 데이터 기록부(5)에 있어서의 특징량의 에러 예측 모델에 대한 감도를 해석함으로써, 에러와 높은 상관을 갖는 특징량을 유저에게 제시할 수 있다.

[0064] 또한, 모델 선정부(14)에서 입력된 데이터에 대하여 성능이 좋은 모델을 선택한다고 함으로써, 데이터 내에 다양한 특징을 갖는 에러 데이터가 혼재하는 경우에도, 관련이 낮은 특징량이 노이즈로서 추출되는 것을 회피하여, 추출되는 특징량의 정밀도를 높일 수 있다.

[0065] 또한, 상술한 모델 선정부(14)에서 2개 이상의 모델이 선정된 경우, 그것들 복수의 모델의 해석 결과를 복합하여 상관이 높은 특징량을 제시한다고 해도 된다.

[0066] 다음으로, 모델 선정부(14)에서 2개의 모델로서 에러 예측 모델 A 및 에러 예측 모델 B가 선택된 경우의 에러 요인 후보의 제시 방법에 대하여 설명한다.

[0067] 도 6은, 본 실시예에 관한 모델 평가부 및 모델 해석부에 있어서의 처리 및 표시예를 도시하는 모식도이다.

[0068] 본 도면에 있어서는, 입력 데이터 기록부(5)의 입력 데이터와 모델 데이터베이스(12)(도 2)의 에러 예측 모델 A(210)가 모델 평가부(13a) 및 모델 해석부(15a)에 입력된다. 또한, 입력 데이터 기록부(5)의 입력 데이터와 모델 데이터베이스(12)(도 2)의 에러 예측 모델 B(211)가 모델 평가부(13b) 및 모델 해석부(15b)에 입력된다. 그리고, 특징량의 에러에 대한 기여도의 그래프(231)가 표시 가능하게 된다.

[0069] 모델 평가부(13a, 13b) 및 모델 해석부(15a, 15b)의 기능은, 도 2와 마찬가지이다. 또한, 도 6에 있어서는, 모델 평가부(13a, 13b) 및 모델 해석부(15a, 15b)를 설명을 위해서 2개씩 도시하고 있지만, 실제로는, 각각 하나의 모델 평가부(13a, 13b) 및 모델 해석부(15a, 15b)에서 순번 혹은 병렬로 처리를 실행하는 것으로 해도 된다.

[0070] 본 도면에 있어서는, 모델 해석부(15a, 15b)에서 계산된 특징량마다의 SHAP값의 총합 ϕ_{SUM}에 대하여, 모델 평가부(13a, 13b)에서 얻어진 모델 평가값으로 보정함으로써 최종적인 기여도 ϕ'_{SUM}을 계산한다. 가장 심플하게는, 예를 들어 모델 평가부(13a, 13b)에서의 모델성 평가 지표가 재현율이었던 경우, 모델의 재현율을 SHAP값 ϕ_{SUM}에 승산함으로써, 최종적인 기여도 ϕ'_{SUM}을 얻는다고 할 수 있다. 각 에러 예측 모델에서 계산된 최종적인 기여도 ϕ'_{SUM}을 값이 큰 순으로 배열하여 표시함으로써, 복수의 모델로부터 복합된 에러와 높은 상관을 갖는 특징량을 랭킹 형식으로 단말기(4)를 통해 유저에게 제시할 수 있다.

[0071] 또한, 이들의 에러 발생에 대하여, 기여도가 높은 랭킹 상위의 특징량에 대하여, 특징량의 값과 에러에 대한 기여도의 관계를 가시화한 예에 대하여 설명한다.

[0072] 도 7은, 본 실시예에 관한 특징량의 값과 파라미터의 에러 발생에 대한 기여도의 관계의 일례를 도시하는 그래프이다. 횡축에 특징량, 종축에 파라미터의 에러 발생에 대한 기여도를 취하고 있다.

[0073] 본 도면에 도시하는 바와 같이, 에러에 대한 기여도의 값에 기초하는 정상 영역 및 위험 영역의 판정이나, 실제의 에러 발생 위치를 단말기(4)의 화면에 표시함으로써, 추출된 특징량의 타당성을 유저가 평가할 수 있다.

[0074] 도 2에 있어서 선정된 에러 요인이 실제와 다른 경우, 유저는, 단말기(4)를 통해 올바른 에러 요인의 지정(수정)을 할 수 있다. 지정이 실시된 경우, 제2 에러 예측 모델 생성부(17)에서 새로운 에러 예측 모델을 생성한다.

[0075] 다음으로, 제2 에러 예측 모델 생성부(17)에 있어서의 모델 생성 수순에 대하여 설명한다.

[0076] 도 8은, 본 실시예에 관한 제2 에러 예측 모델 생성 수순을 도시하는 흐름도이다.

[0077] 본 도면에 있어서는, 유저에 의한 원인 파라미터의 수정이 있는 경우(스텝 S800)에, 스텝 S201로 이행한다.

[0078] 스텝 S201에 있어서는, 에러 예측 모델의 생성에 필요한 학습 데이터의 선정(추출)을 한다. 선정 방법은, 도 3의 스텝 S101과는 다르다. 도 8의 스텝 S201에 있어서는, 에러 예측 모델에서 검출된 에러에 대하여 동일한 레시피 또는 유사한 레시피를 포함하는 데이터를 추출한다.

[0079] 본 도면에 있어서의 이후의 S102로부터 S105까지의 처리는, 도 3과 마찬가지이다.

[0080] 이어서, 스텝 S202에 있어서, 에러 예측 모델을 해석하고, 스텝 S201에서 추출된 학습 데이터 내의 특징량에 있어서의 에러 판정에 대한 기여도를 수치화한다. 이것은, 도 2에 있어서의 모델 해석부(15)와 마찬가지의 처리이다.

[0081] 스텝 S203에 있어서, 미리 정해진 순위 내의 기여도가 높은 특징량 중에, 유저로부터 지정된 에러 요인이 포함되어 있지 않으면 S102로부터 마찬가지의 처리를 반복하고, 포함되어 있으면 신규 에러 예측 모델의 생성 완료로서 모델 데이터베이스(12)에 보존한다(스텝 S300).

[0082] 이와 같이, 기계 학습으로 정답에 닿을 수 없는 경우, 사람 손으로 정답을 가르침으로써 학습 모델을 고도화할 수 있다.

[0083] **실시예 2**

[0084] 도 9는, 실시예 2에 관한 데이터 분류부를 구비한 에러 요인 추정 장치를 도시하는 구성도이다.

[0085] 본 실시예(본 도면)와 실시예 1의 차이는, 입력 데이터 기록부(5)로부터 보내진 데이터를 분류하는 데이터 분류부(18)를 에러 요인 추정 장치(3)에 마련한 것이다. 다른 구성은, 도 2와 마찬가지이다.

[0086] 실시예 1에 있어서는, 동일한 레시피에서 발생하는 에러는 동일한 요인라는 상정 하에, 동일한 레시피 또는 유

사한 레시피 데이터를 사용하여, 모델 평가·모델 생성을 실행하고 있다.

[0087] 이에 비해, 본 실시예(도 9)에 있어서는, 데이터 분류부(18)에서 에러 요인마다 에러 데이터를 분류하고, 분류된 에러 데이터마다 모델 평가·모델 생성을 실행한다.

[0088] 다음으로, 데이터 분류부(18)의 구성에 대하여 설명한다.

[0089] 도 10은, 도 9의 데이터 분류부의 상세를 도시하는 구성도이다.

[0090] 도 10에 있어서는, 데이터 분류부(18)는, 에러 예측 모델 생성부(19)와, 모델 해석부(115)와, 에러 요인 클러스터링부(20)와, 데이터 분할부(21)와, 분할 데이터 기록부(122)를 구비하고 있다.

[0091] 에러 예측 모델 생성부(19)는, 입력 데이터 기록부(5)로부터 보내진 입력 데이터에 포함되는 에러 발생의 유무의 판정에 사용되는 학습 모델인 에러 예측 모델을 생성한다. 에러 예측 모델 생성부(19)에 있어서의 공정에 있어서는, 도 8의 스텝 S103과 마찬가지의 것을 사용할 수 있다.

[0092] 모델 해석부(115)는, 에러 예측 모델 생성부(19)에서 생성된 모델의 판정 결과에 각 특징량이 어느 정도 기여했는지를, 예를 들어 SHAP값을 사용하여 계산한다.

[0093] 에러 요인 클러스터링부(20)는, 모델 해석부(115)에서 계산된 SHAP값으로 대표되는 각 특징량의 에러에 대한 기여도에 대해서, 교사 없이 학습을 적용하여 분류한다.

[0094] 데이터 분할부(21)는, 분류된 에러와 정상 데이터로 분할한다. 분할된 데이터는, 분할 데이터 기록부(122)에 보존된다.

[0095] 다음으로, 이 분류의 사고 방식에 대하여 설명한다.

[0096] 도 11은, 특징량의 값과 파라미터의 에러 발생에 대한 기여도의 관계를 사용하여 에러 데이터를 분류한 상태를 도시하는 그래프이다. 횡축에 특징량, 종축에 파라미터의 에러 발생에 대한 기여도를 취하고 있다.

[0097] 본 도면에 있어서는, 다른 에러 요인에 대응하는 분류 1, 2를 각각의 영역으로 분할하고, 각각에 있어서 에러 발생으로 된 데이터(에러 데이터)를 분리한다. 이것은, 요인이 다른 에러에서는, 그것을 판별하기 위한 모델 내부의 분기식도 다르기 때문이고, 에러 판정에 대한 기여도를 나타내는 SHAP값은, 에러 요인에 관련되는 특징량마다 그 존재 범위가 다른 것이 기대되기 때문이다. 에러 데이터는, 이 SHAP값에 따라서 분리한다.

[0098] 이와 같이 하여 에러 요인마다 분리된 에러 데이터에 대하여, 도 9의 제1 에러 예측 모델 생성부(16)나 제2 에러 예측 모델 생성부(17)에서 에러 예측 모델을 생성하고, 모델 해석부(15)에서 모델을 해석함으로써, 입력 데이터 기록부(5)의 입력 데이터 내에 모델 데이터베이스(12) 내의 에러 예측 모델에서 대응할 수 없는 에러 요인이 복수 존재하는 경우에도, 관련이 낮은 특징량이 노이즈로서 추출되는 것을 회피하여, 추출되는 특징량의 정밀도를 높일 수 있다. 여기서, 데이터 분류부(18)에서 분할된 에러 데이터를 사용하여 에러 예측 모델을 생성할 때에는, 대상의 에러 데이터와 동일한 레시피 또는 유사한 레시피인 정상 데이터를 합쳐서 학습 데이터로서 사용할 수 있다.

[0099] 또한, 분류된 에러 데이터마다 다른 라벨 부여를 하여, 이 라벨과 함께 에러 예측 모델을 생성하여 모델 데이터베이스(12)에 보존하는 것으로 해도 된다. 이 라벨 부여는, 다른 인덱스를 순서대로 자동으로 할당한다고 해도 되고, 유저가 에러 요인을 라벨 부여하는 것으로 해도 된다. 이 에러 요인의 라벨 부여는, 분할된 데이터 단위로 실시하면 되기 때문에, 종래 방법인 발생 에러 1건 단위에 대한 라벨 구여와 비교하여, 처리 수를 대폭으로 삭감할 수 있다. 이 경우, 도 9에 있어서 모델 평가부(13)가 불필요하게 되고, 모델 선정부(14)의 기능이 에러 데이터의 라벨과 에러 예측 모델의 라벨이 일치하는 것을 선택하는 동작으로 치환된다.

[0100] **실시예 3**

[0101] 실시예 1의 에러 요인 추정 장치에 있어서, 도 2의 모델 데이터베이스(12) 내에 충분한 에러 예측 모델의 베리에이션이 보존된 때에는, 새로운 에러 예측 모델은 생성되지 않고, 입력된 에러 데이터에 대하여 에러 요인의 추정만을 실시하는 운용으로 해도 된다.

[0102] 이 경우에 있어서의 에러 요인 추정 장치의 구성에 대해서, 실시예 3으로서 설명한다.

[0103] 도 12는, 본 실시예의 에러 요인 추정 장치를 도시하는 구성도이다.

[0104] 본 도면에 있어서는, 데이터베이스(22)로부터, 에러 데이터로서, 에러가 발생한 측정점에 있어서의 레시피·계

측 결과나, 그것과 동일한 웨이퍼이며 동일한 레시피인 것에 있어서의 레시피·계측 결과의 데이터 세트가 입력된다.

[0105] 특징량 생성부(11)에서, 이들의 데이터에 대하여 기계 학습 모델에 적합한 특징량을 추출하고, 모델 입력 에러 데이터 기록부(23)에 에러 데이터를 출력한다.

[0106] 이후의 처리는, 도 2와 마찬가지이고, 모델 입력 에러 데이터 기록부(23)의 에러 예측에 대하여 평가값이 높은 에러 예측 모델을 하나 이상 모델 데이터베이스(12) 내로부터 선택하고, 그 모델의 해석 결과로서 에러 예측에 기여도가 큰 특징량을, 단말기(4)를 통해 유저에게 제시한다. 이 경우, 모델 선정부(14)는, 불필요하게 된다.

[0107] 이렇게 에러 요인 추정 장치(3)의 기능을 발생한 에러의 요인 추정으로 한정함으로써, 필요해지는 입력 데이터도 최소한으로 할 수 있다.

[0108] 도 13은, 에러 요인 추정 장치의 변형예를 도시하는 구성도이다.

[0109] 본 도면에 있어서는, 에러 요인 추정 장치(3)는, 에러 요인 라벨 취득부(24)와, 에러 요인 라벨 데이터베이스(25)(에러 요인 라벨 DB)를 구비하고 있다. 이에 의해, 모델 해석부(15)에서 추출된 특징량에 기초하여 에러 요인 후보를 취득할 수 있다.

[0110] 에러 요인 라벨 데이터베이스(25)에는, 각 특징량이나 그 조합에 대응하는 에러 요인의 관계를 보존해 둔다. 이 경우, 특징량에 대한 에러 요인의 라벨 부여가 미리 필요해지지만, 종래 방법인 발생 에러 1건 단위에 대한 라벨 부여와 비교하여, 필요한 처리 수를 대폭으로 삭감할 수 있다.

[0111] 에러 요인 라벨 취득부(24)에 있어서는, 모델 해석부(15)에서 얻어진 랭킹 상위의 특징량에 대하여 에러 요인 라벨 데이터베이스(25) 내의 라벨 관계를 사용하여 대응하는 에러 요인을 부여한다.

[0112] 이 라벨 부여된 에러 요인을, 단말기(4)를 통해 유저에게 제시한다. 이때, 모델 해석부(15)에서 계산된 각 특징량의 기여도의 크기를, 대응하는 에러 요인의 정확도로서 환산하여 표시한다고 해도 된다.

[0113] 다음으로, 이 표시의 예에 대하여 설명한다.

[0114] 도 14는, 본 실시예에 관한 유저에 대한 에러 요인 후보의 표시예를 도시한 것이다.

[0115] 도면 중 우측의 그래프는, 모델 해석부(15)에서 계산된 각 특징량의 기여도의 크기를 나타내고 있다. 도면 중 좌측에는, 우측에 그래프에 나타내는 데이터를 기초로 환산한 에러 요인의 정확도를 나타내고 있다.

[0116] 이와 같이, 에러 예측 모델을 해석해서 에러 판정에 기여한 특징량을 계산하고, 각 특징량이나 그 조합에 대응한 에러 요인이 라벨 부여됨으로써, 발생한 에러와 상관이 높은 특징량이나 대응하는 에러 요인 후보를 유저에게 제시할 수 있다.

[0117] (기타)

[0118] 본 발명은 전술한 실시예에 한정되는 것은 아니고, 다양한 변형예가 포함된다. 상기의 실시예는, 본 발명을 이해하기 쉽게 설명하기 위하여 상세하게 설명한 것이고, 반드시 설명한 모든 구성을 구비하는 것에 한정되는 것은 아니다. 예를 들어, 실시예 3의 도 13에 도시하는 에러 요인 라벨 취득부(24) 및 에러 요인 라벨 데이터베이스(25)는, 실시예 1의 도 2 및 실시예 2의 도 9에도 조합할 수 있고, 각각 특징량이나, 특징량의 조합에 관련지어진 에러 요인을 유저에게 나타낼 수 있다.

[0119] 전술한 실시예에 있어서는, 반도체 검사 장치의 에러 요인을 추정하는 예를 설명했지만, 장치의 동작을 규정하는 파라미터 및 그 파라미터를 채용했을 때에 에러가 발생하는지의 여부에 관한 예측 모델을 생성시킴으로써, 반도체 검사 장치 이외의 장치에 대해서도 본 발명을 적용할 수 있다.

[0120] 또한, 전술한 실시예에 있어서는, 에러 예측 모델에 대한 특징량의 기여도에 대하여 SHAP값을 사용하여 수치화하는 예를 나타내고 있지만, Feature Importance 등, 다른 평가값을 적용할 수 있다.

부호의 설명

[0121] 1: 반도체 검사 장치

2, 22: 데이터베이스

3: 에러 요인 추정 장치

4: 단말기

5: 입력 데이터 기록부

11: 특징량 생성부

12: 모델 데이터베이스

13, 13a, 13b: 모델 평가부

14: 모델 선정부

15, 15a, 15b, 115: 모델 해석부

16: 제1 에러 예측 모델 생성부

17: 제2 에러 예측 모델 생성부

18: 데이터 분류부

19: 에러 예측 모델 생성부

20: 에러 요인 클러스터링부

21: 데이터 분할부

23: 모델 입력 에러 데이터 기록부

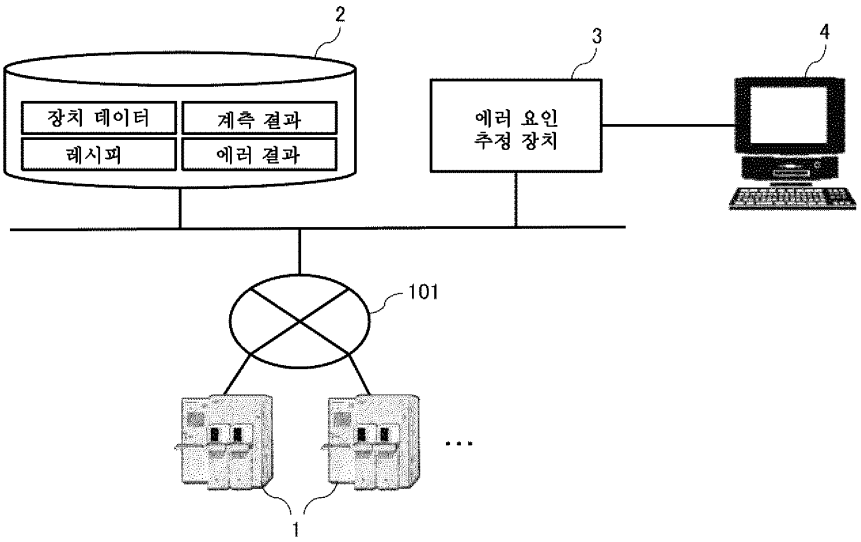
24: 에러 요인 라벨 취득부

25: 에러 요인 라벨 데이터베이스

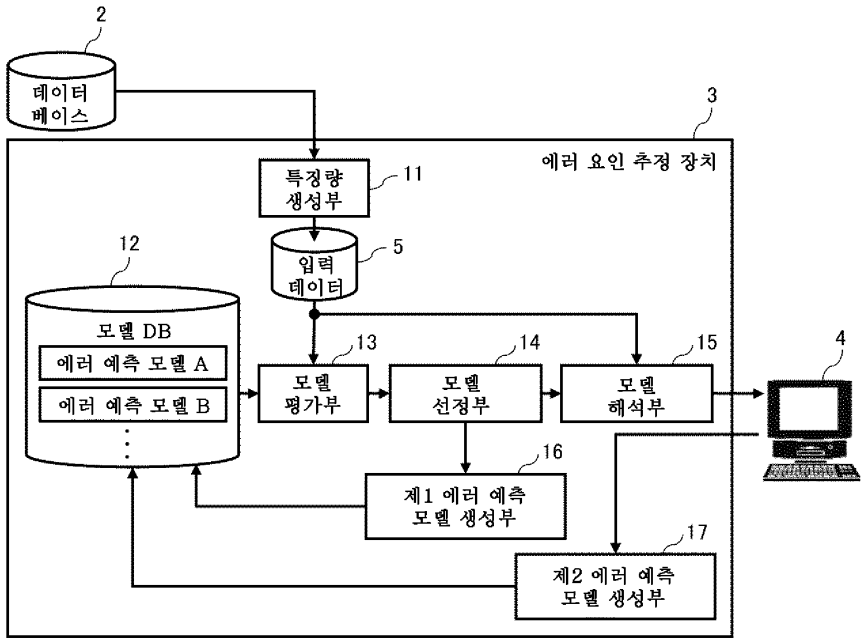
101: 네트워크.

도면

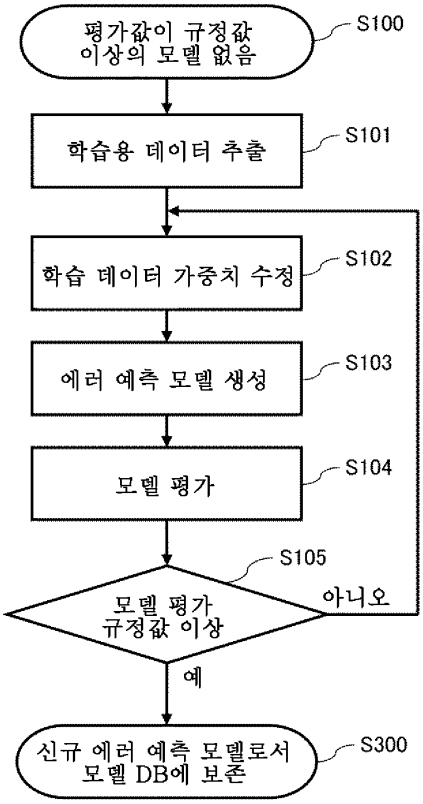
도면1



도면2


2
데이터
베이스
3
에러 요인 추정 장치
특징량
생성부
11
입력
데이터
5
12
모델 DB
에러 예측 모델 A
에러 예측 모델 B
13
모델
평가부
14
모델
선정부
15
모델
해석부
4
16
제1 에러 예측
모델 생성부
17
제2 에러 예측
모델 생성부

도면3

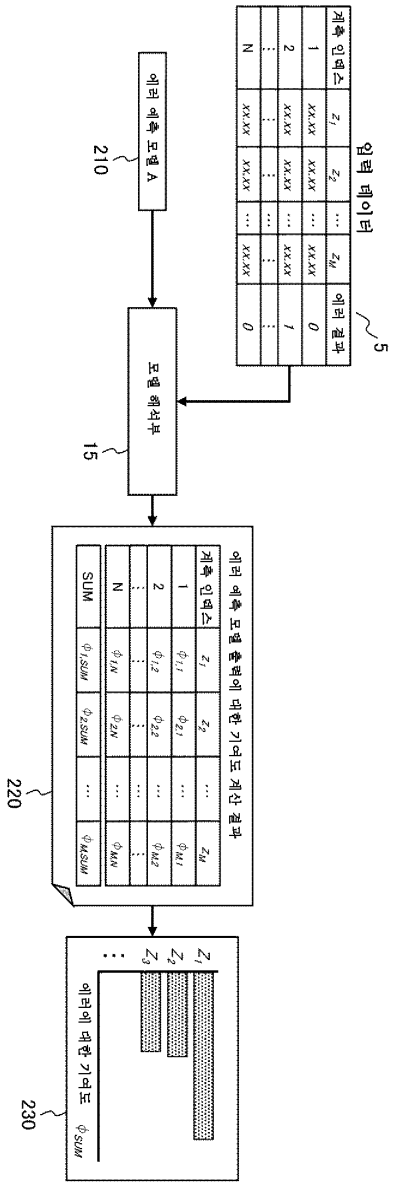

평가값이 규정값
이상의 모델 없음
S100
학습용 데이터 추출
S101
학습 데이터 가중치 수정
S102
에러 예측 모델 생성
S103
모델 평가
S104
S105
모델 평가
규정값 이상
아니오
예
신규 에러 예측 모델로서
모델 DB에 보존
S300

도면4

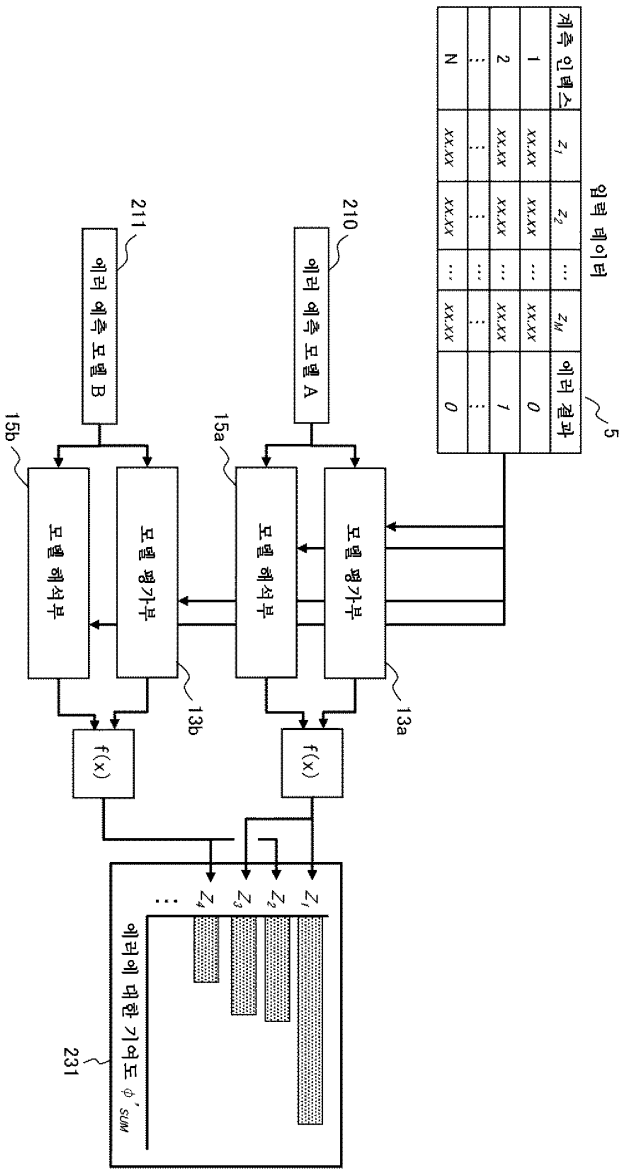
측정 인덱스	z_1	z_2	…	z_M	동작 공정	에러 결과
1	xx.xx	xx.xx	…	xx.xx	1	0
2	xx.xx	xx.xx	…	xx.xx	2	1
3	xx.xx	xx.xx	…	xx.xx	2	0
4	xx.xx	xx.xx	…	xx.xx	3	0
…						

학습 데이터

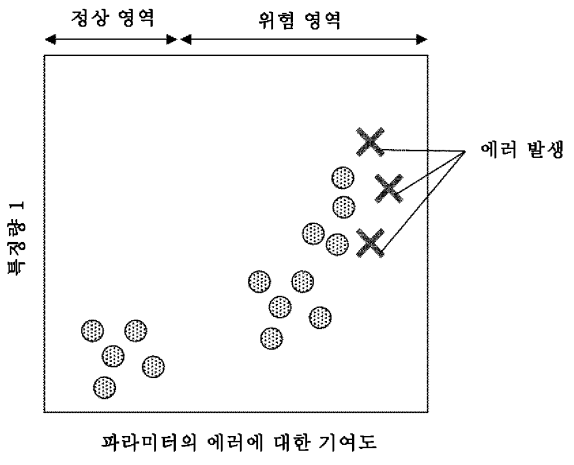
도면5



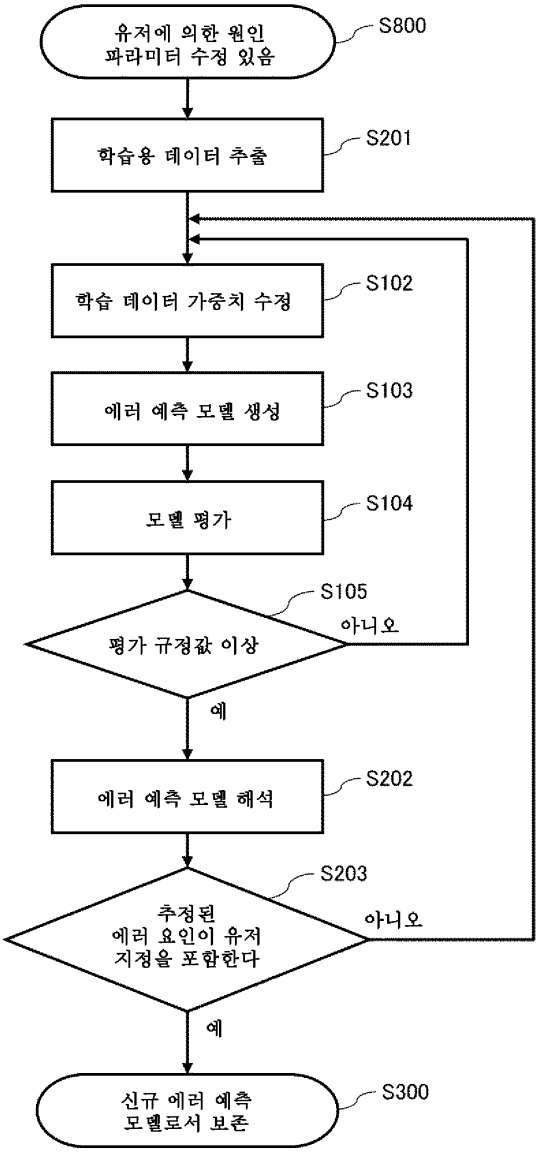
도면6


입력 데이터
계측 인덱스
에러 결과
5
210
에러 예측 모델 A
211
에러 예측 모델 B
15a
15b
모델 해석부
모델 평가부
13a
13b
f(x)
에러에 대한 기여도
231

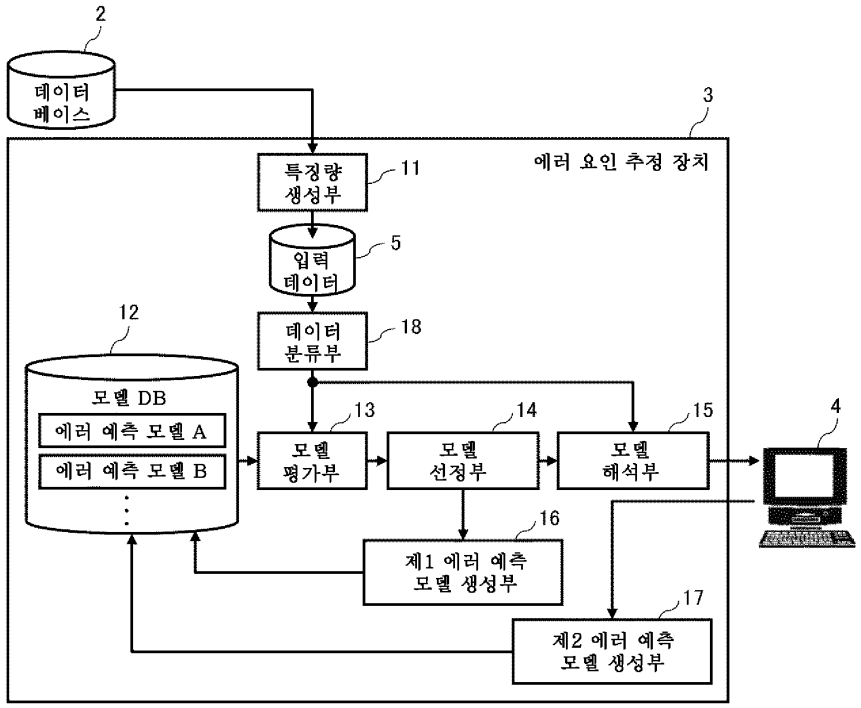
도면7


정상 영역
위험 영역
에러 발생
특징량 1
파라미터의 에러에 대한 기여도

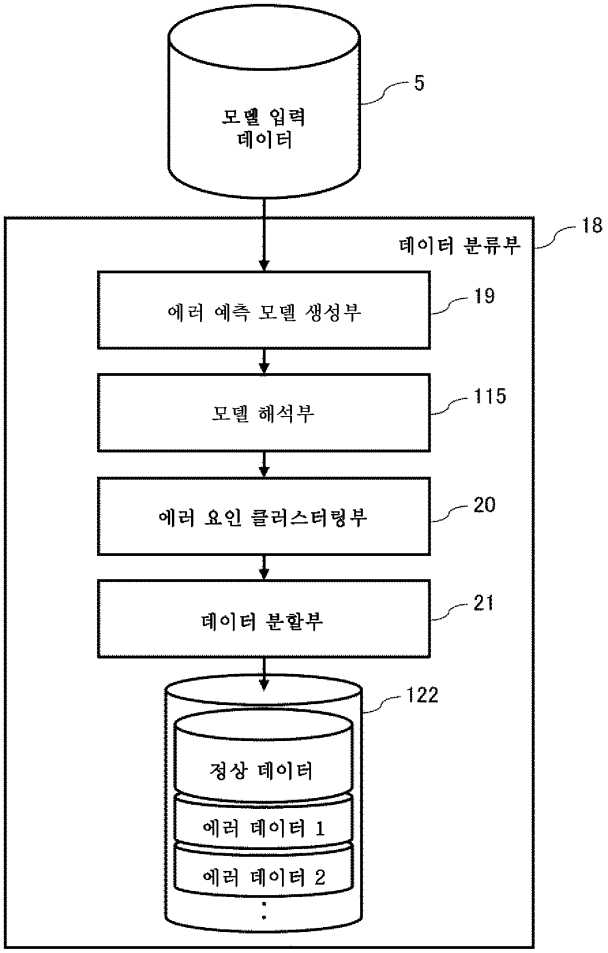
도면8


유저에 의한 원인
파라미터 수정 있음
S800
학습용 데이터 추출
S201
학습 데이터 가중치 수정
S102
에러 예측 모델 생성
S103
모델 평가
S104
S105
평가 규정값 이상
아니오
예
에러 예측 모델 해석
S202
S203
추정된
에러 요인이 유저
지정을 포함한다
아니오
예
신규 에러 예측
모델로서 보존
S300

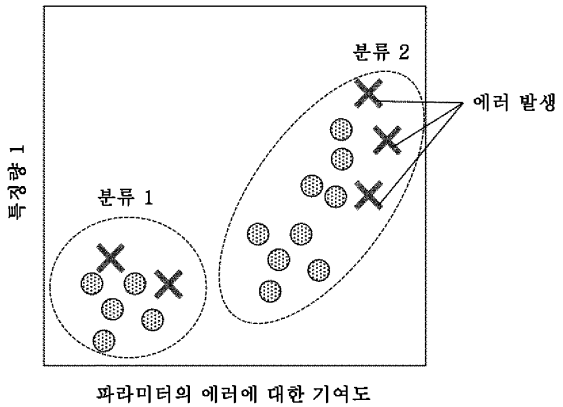
도면9


2
데이터
베이스
3
에러 요인 추정 장치
특징량
생성부
11
입력
데이터
5
데이터
분류부
18
12
모델 DB
에러 예측 모델 A
에러 예측 모델 B
13
모델
평가부
14
모델
선정부
15
모델
해석부
4
16
제1 에러 예측
모델 생성부
17
제2 에러 예측
모델 생성부

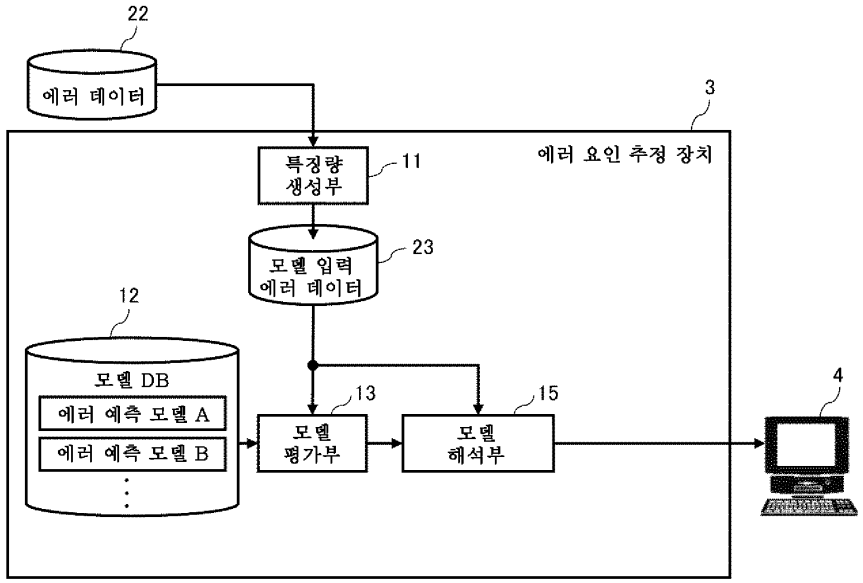
도면10


모델 입력 데이터
5
데이터 분류부
18
에러 예측 모델 생성부
19
모델 해석부
115
에러 요인 클러스터링부
20
데이터 분할부
21
122
정상 데이터
에러 데이터 1
에러 데이터 2

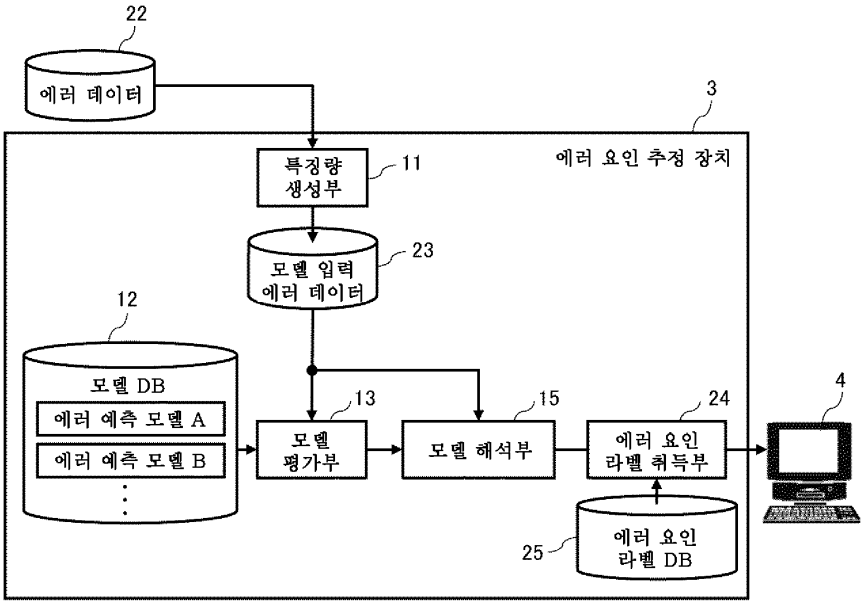
도면11


분류 2
에러 발생
분류 1
특징량 1
파라미터의 에러에 대한 기여도

도면12



도면13



도면14

