



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년03월08일
(11) 등록번호 10-1955305
(24) 등록일자 2019년02월28일

- | | |
|--|---|
| <p>(51) 국제특허분류(Int. Cl.)
G01D 3/08 (2006.01) G05B 23/02 (2006.01)</p> <p>(52) CPC특허분류
G01D 3/08 (2013.01)
G05B 23/024 (2013.01)</p> <p>(21) 출원번호 10-2017-7009440</p> <p>(22) 출원일자(국제) 2015년09월03일
심사청구일자 2017년04월07일</p> <p>(85) 번역문제출일자 2017년04월06일</p> <p>(65) 공개번호 10-2017-0045354</p> <p>(43) 공개일자 2017년04월26일</p> <p>(86) 국제출원번호 PCT/US2015/048265</p> <p>(87) 국제공개번호 WO 2016/040082
국제공개일자 2016년03월17일</p> <p>(30) 우선권주장
62/048,588 2014년09월10일 미국(US)</p> <p>(56) 선행기술조사문헌
US20070124113 A1*
*는 심사관에 의하여 인용된 문헌</p> | <p>(73) 특허권자
지멘스 에너지, 인코포레이티드
미국 플로리다주 올랜도 알라파야 트레일 4400 (우: 32826-2399)</p> <p>(72) 발명자
고, 시용 테
미국 02139 매사추세츠 캄브리지 메모리얼 드라이브 550 아파트먼트 10이-4
유안, 차오
미국 08536 뉴저지 플레인스버러 매리언 드라이브 36
(뒷면에 계속)</p> <p>(74) 대리인
특허법인 남앤남</p> |
|--|---|

전체 청구항 수 : 총 10 항

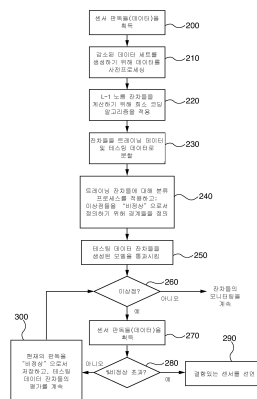
심사관 : 김려원

(54) 발명의 명칭 **희소 코딩 방법을 활용한 가스 터빈 센서 고장 검출**

(57) 요약

가스 터빈들을 모니터링하는 데 사용되는 센서들의 고장들을 인식(및/또는 예측)하기 위한 방법 및 시스템은, 희소 코딩 프로세스를 수집된 센서 판독들에 적용하고 그리고 희소 코딩 프로세스로부터의 L-1 노름 잔차들이 잠재적 센서 문제를 표시하는 것으로 정의한다. 잔차 센서 판독들의 그룹의 추가의 평가는 그룹을 카테고리화하기 위해 그리고 잡음있는 데이터보다는 결합있는 센서와 더 연관될 가능성이 있는 것으로 고려될 상당한 이상점들(비정상 데이터)이 존재하는지를 결정하기 위해 수행된다. 현재의 비정상 결과를 이전 결과들의 세트와 비교하는 평가에 시간 컴포넌트가 도입되며, 상당한 수의 이전 판독들이 또한 비정상 값을 갖는 경우 결합있는 센서 결정을 한다. 시간 컴포넌트를 고려함으로써, 허위 포지티브들의 수가 감소된다.

대표도



(72) 발명자

차크라보르티, 아미트

미국 08520 뉴저지 이스트 원저 인버네스 레인 2

에반스, 매튜

미국 32789 플로리다 윈터 파크 밴코리 로드 1826

명세서

청구범위

청구항 1

적어도 하나의 센서(sensor)(7)로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장(gas turbine sensor failure)을 검출하기 위한 방법으로서,

정상 동작 상태(normal operating condition)들에 대해 알려진 센서 판독들과 연관된 값들을 정의하는 기저 벡터(basis vector)들의 딕셔너리(dictionary)를 생성하는 단계(520);

L-1 노름 잔차 데이터(L-1 norm residual data)의 세트(set)를 식별하기 위해, 생성된 딕셔너리를 사용하여, 희소 코딩 프로세스(sparse coding process)를 센서 판독 데이터의 세트에 적용하는 단계(530, 600);

최대-값의 L-1 노름 잔차 데이터(largest-valued L-1 norm residual data)의 미리 결정된 서브세트(subset)를 비정상 센서 판독(abnormal sensor reading)들로 카테고리화(categorize)하기 위해 상기 L-1 노름 잔차 데이터를 평가하는 단계(610);

상기 비정상 센서 판독들을 복수의 이전(prior) 센서 판독들과 비교하고 그리고 상기 이전 센서 판독들 중 미리 정의된 개수의 이전 센서 판독들이 또한 비정상 센서 판독들로서 지정되는 경우, 지정된 비정상 센서 판독들을 센서 고장과 연관된 것으로 정의하는 단계(620, 710);

특정 센서를 식별하는 센서 고장 신호를 가스 터빈 직원(gas turbine personnel)에게 송신하는 단계; 및

고장난 센서를 서비스(service)로부터 제거하고 그리고 상기 고장난 센서를 수리 또는 교체하는 단계(660, 770)를 포함하는,

적어도 하나의 센서(7)로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법.

청구항 2

제 1 항에 있어서,

상기 딕셔너리를 생성하는 단계에서,

상기 센서 판독 데이터를 복수의 클러스터(cluster)들로 조직화하는 단계(522),

k-평균 알고리즘(k-means algorithm)을 사용하여 각각의 클러스터의 센터(center)를 결정하는 단계(524), 및

각각의 클러스터의 결정된 센터를 이용하여 상기 딕셔너리를 초기화하는 단계(526)가 수행되는,

적어도 하나의 센서(7)로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법.

청구항 3

제 1 항에 있어서,

상기 희소 코딩 프로세스는 아래의 관계식,

$$\min_{E, X} \|Y - DX - E\|_F^2 + \alpha \|X\|_{1,2} + \beta \|E\|_{1,1},$$

을 수행하고,

여기서 D는 생성된 딕셔너리이고, X는 복수의 센서 판독들이고, E는 L-1 노름 잔차들이고, 그리고 α 및 β 는 작은 포지티브 스칼라(small positive scalar)들인,

적어도 하나의 센서(7)로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법.

청구항 4

제 3 항에 있어서,

최소화 함수(minimization function)는 아래의 반복적 단계들,

a.
$$X^{k+1} = \arg \min_X L(X, Z^k, E^k, M^k, \mu^k) = (D^T D + \mu^k I)^{-1} (D^T (Y - E^k) + M^k + \mu^k Z^k)$$
 - 여기서 L은 증강된-라그랑지안(Augmented-Lagrangian)이고 그리고

$$L(X, Z, E, M, \mu) = \frac{1}{2} \|Y - DX - E\|_F^2 + \alpha \|Z\|_{1,2} + \beta \|Z\|_{1,1} + \langle M, Z - X \rangle + \frac{\mu}{2} \|Z - X\|_F^2$$
 로서 정의됨 -,

b.
$$Z^{k+1} = \arg \min_Z L(X^{k+1}, Z, E^k, M^k, \mu^k) = \arg \min_Z \frac{1}{2} \|P - Z\|_F^2 + \gamma \|Z\|_{1,2}$$
 - 여기서

$$P = X^{k+1} - \frac{1}{\mu^k}$$
 이고, 행 수축 연산자(row shrinkage operator), 즉,

$$Z^{k+1}(j,:) = \begin{cases} \frac{\|P(j,:)\|_2 - \gamma}{\|P(j,:)\|_2} P(j,:) & \gamma < \|P(j,:)\|_2 \text{ 인 경우} \\ 0 & \text{그렇지 않은 경우} \end{cases}$$
 를 사용하는 경우 -,

c.
$$E^{k+1} = \arg \min_E L(X^{k+1}, Z^{k+1}, E, M^k, \mu^k) = \arg \min_Z \frac{1}{2} \|Q - E\|_F^2 + \beta \|E\|_{1,1}$$
 - 여기서

$$Q = Y - DX^{k+1}$$
 및
$$E^{k+1} = S_\beta(Q)$$
 이고, 여기서
$$S_\beta(a) = \begin{cases} a - \beta & a > \beta \text{ 인 경우} \\ 0 & |a| \leq \beta \text{ 인 경우} \\ a + \beta & a < -\beta \text{ 인 경우} \end{cases}$$
 임 -,

d.
$$M^{k+1} = M^k + \mu^k (Z^{k+1} - X^{k+1})$$

e.
$$\mu^{k+1} = \rho \mu^k$$

f.
$$k = k + 1$$

g.
$$\frac{\|Z^k - X^k\|_F^2}{\|X^k\|_F^2} < \epsilon$$
 인 경우 중지됨 - 여기서 E^k 는 L-1 정상 잔차 데이터의 세트로서 정의되고, Z는 보조 변수이고, M은 라그랑주 승수(Lagrange multiplier)이고, μ 는 페널티 계수(penalty coefficient)이고, ρ 는 μ 를 업데이트(updating)하기 위한 파라미터(parameter)이고, 그리고 ϵ 는 알고리즘을 종결시키기 위해 Z 내지 X의 근사(closeness)를 측정하는 데 사용되는 에러 허용한계(error tolerance)임 -

을 수행하는,

적어도 하나의 센서(7)로부터 수집된 관독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법.

청구항 5

제 1 항에 있어서,

상기 평가를 수행하는 단계에서,

일-클래스 지원 벡터 머신 모델(one-class Support Vector Machine model)이 생성되고(550), 그리고 각각의 잔차는 상기 클래스 내에 있는 것 또는 상기 클래스 외측에 있는 것으로 분류되고, 상기 클래스 외측의 L-1 노름 잔차들은 비정상 센서 관독들로서 정의되고,

상기 일-클래스 SVM 모델을 생성하기 전에 상기 데이터 세트의 크기를 감소시키기 위해, 상기 L-1 노름 잔차 데이터의 세트가 샘플링되는(sampled)(540),

적어도 하나의 센서(7)로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법.

청구항 6

제 5 항에 있어서,

상기 비교를 수행하는 단계에서,

복수의 이전 시간 단계 데이터(prior time step data)가 조사되고(reviewed)(640), 그리고 상기 복수의 이전 시간 단계 데이터의 p%가 또한 비정상 센서 판독으로서 지정되는 경우, 센서 고장 결정이 이루어지는(650),

적어도 하나의 센서(7)로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법.

청구항 7

제 6 항에 있어서,

상기 p의 값은 적어도 50과 동일한,

적어도 하나의 센서(7)로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법.

청구항 8

제 1 항에 있어서,

상기 L-1 노름 잔차 데이터는, 임계치 퍼센티지(threshold percentage)를 정의하고(725) 그리고 정의된 임계치를 초과하는 모든 값들을 비정상 센서 판독 데이터로서 정의함으로써(740) 평가되는,

적어도 하나의 센서(7)로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법.

청구항 9

제 8 항에 있어서,

상기 비교를 수행하는 단계(760)에서,

복수의 이전 시간 단계 데이터가 조사되고, 그리고 상기 복수의 이전 시간 단계 데이터의 q%가 또한 비정상 센서 판독으로서 지정되는 경우, 센서 고장 결정이 이루어지는(770),

적어도 하나의 센서(7)로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법.

청구항 10

가스 터빈 센서(7)의 고장을 검출하기 위한 시스템(system)으로서,

센서 판독들의 데이터베이스(database)(12);

상기 센서 판독들의 데이터베이스(12)와 통신하는 센서 모니터링 시스템 컴포넌트(sensor monitoring system component)(10) - 상기 센서 모니터링 시스템 컴포넌트(10)는 프로그램 저장 디바이스(program storage device)(16) 및 프로세서(processor)(18)를 포함하고, 상기 프로그램 저장 디바이스(16)는 방법 단계들을 수행하기 위해 상기 프로세서(18)에 의해 실행가능한 프로그램 명령들의 세트를 고정형의 유형적 매체(fixed tangible medium)로 포함하고, 상기 방법 단계들은,

정상 동작 상태들에 대해 알려진 센서 판독들과 연관된 값들을 정의하는 기저 벡터들의 디서너리를 생성하는 단계(520),

L-1 노름 잔차 데이터의 세트를 식별하기 위해, 생성된 디서너리를 사용하여, 희소 코딩 프로세스를 센서 판독 데이터의 세트에 적용하는 단계(530, 600),

최대-값의 L-1 노름 잔차 데이터의 미리 결정된 서브세트를 비정상 센서 판독들로 카테고리화하기 위해

상기 L-1 노름 잔차 데이터를 평가하는 단계(610),

상기 비정상 센서 판독들을 복수의 이전 센서 판독들과 비교하고 그리고 상기 이전 센서 판독들 중 미리 정의된 개수의 이전 센서 판독들이 또한 비정상 센서 판독들로서 지정되는 경우, 지정된 비정상 센서 판독들을 센서 고장과 연관된 것으로 정의하는 단계(620, 710), 및

수리 또는 교체되도록 하기 위해, 고장난 센서를 식별하는 정보를 포함하는 센서 고장 신호를 가스 터빈 직원에게 송신하는 단계(660, 770)임 -;

비정상(abnormal)으로서 분류된 센서 판독들을 저장하기 위한 모니터링 시스템 데이터베이스(20)를 포함하는, 가스 터빈 센서(7)의 고장을 검출하기 위한 시스템.

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

발명의 설명

기술 분야

[0001] 관련 출원들에 대한 상호-참조

[0002] 본 출원은, 2014년 9월 10일 출원된 미국 가출원 번호 제 62/048,588호를 우선권으로 주장하며, 이 미국 가출원은 인용에 의해 본원에 포함된다.

[0003] 기술 분야

[0004] 본 발명은, 전력 생성 네트워크(power generation network)들에서 활용되는 가스 터빈(gas turbine)들의 관리

감독(supervision), 및 특정 센서(sensor)의 고장(failure)을 예측하는 능력을 포함하여 터빈의 성능을 모니터링(monitor)하기 위해 사용되는 센서들 중 임의의 센서의 고장을 검출하는 능력을 개선하기 위한 방법론에 관한 것이다.

배경 기술

- [0005] 화석 연료들의 효율적인 사용은 안정적인 전력 네트워크를 유지하는 데 중요하다. 이러한 유형의 연료를 전기 에너지(electrical energy)로 변환하는 특히 효율적인 수단은 가스 터빈이다. 가스 터빈 컴포넌트(component)들은 매우 고온의 환경에서 그리고 다양한 부하 상태들 하에서 동작한다. 열 피로 및 마모로 인한 부분들의 열화는 현실적인 관심사이다. 효율적인 동작을 계속해서 보장하기 위하여 마모를 검출 및 제어하기 위해서뿐만 아니라 필요에 따라서는 마모된 부분들을 수리 또는 교체하기 위해서 유지보수가 수행된다.
- [0006] 가스 터빈의 성능은 통상적으로, 가스 터빈의 동작의 다양한 양상들을 평가하기 위해 다양한 상이한 센서들(즉, 전력 센서들, 온도 센서들, 압력 센서들 등)을 사용함으로써 모니터링된다(monitored). 불행하게도, 센서 관독(sensor reading)들 그 자체는 상대적으로 잡음이 있는 경향이 있어서, 때때로 센서가 적절하게 동작하는지를 알기 어렵다.
- [0007] 결함있는 센서(faulty sensor)들을 검출하는 문제를 연구하기 위해 사용된 다양한 접근법들이 종래기술에 존재한다. 예컨대, 일 접근법은 센서 관독들에서 3개의 특정 유형들의 에러(error)들, 즉, 짧은 결함(short fault)들, 지속적인 결함(constant fault)들 및 잡음 결함(noise fault)들을 모니터링한다. 센서 관독들에서 "결함있는 센서" 시그니처(signature)들을 탐색하기 위해, 주요 컴포넌트 분석(PCA; principal component analysis)을 활용하는 전통적인 통계적 방법론들이 수행되었다.
- [0008] 가스 터빈 센서들의 결함 검출을 수행하기 위한 이들 접근법들 및 다른 접근법들은 임계치-기반 접근법을 사용한다. 즉, 특정 센서 관독이 주어진 임계치를 초과하는 경우, 센서는 "결함있음(faulty)"으로 선언된다. 작업 가능한 동안, 이 접근법은 센서의 작업 환경의 시간 컴포넌트를 고려하지 않는다. 부가적으로, 일부 센서 관독들에는 큰 잡음 컴포넌트가 존재할 수 있기 때문에, 임계치 접근법은 잡음있는 데이터(noisy data)와 결함있는 센서 간의 차이를 판별하는 데 어려움을 겪을 수 있다. 결과적으로, 이 임계치-기반 접근법은 많은 수의 허위 포지티브(false positive)들을 산출할 수 있는데, 이들은 센서들의 불필요한 검사들로 옮겨져서, 가스 터빈의 성능을 방해할 수 있다.

발명의 내용

- [0009] 종래 기술의 남아있는 필요성들은, 전력 생성 네트워크들에서 활용되는 가스 터빈들의 관리감독 및 특정 센서의 고장을 예측하는 능력을 포함하여 센서 고장을 정확하게 검출하는 능력을 개선하기 위한 방법론에 관한 본 발명에 의해 처리된다.
- [0010] 본 발명에 따르면, 센서 관독들을 분석하기 위해 희소 코딩 프로세스(sparse coding process)가 사용된다. 희소 코딩 접근법은 이 태스크(task)에 아주 적합한 것으로 고려되는데, 그 이유는 정상 센서 관독(normal sensor reading)들의 방대한 집합 가운데서 최소 개수의 "결함있는 센서" 관독들을 식별하는 것이 그 목적이기 때문이다. 희소 코딩 프로세스로부터의 잔차(residual)들은 잠재적으로 비정상 센서 관독(abnormal sensor reading)들로서 정의되며, 시간 팩터(time factor)를 분석에 도입하기 위해 그리고 잔차를 "잡음있는 데이터" 또는 "결함있는 센서"로서 정의하기 위한 결정을 하기 위해 추가로 프로세싱된다(processed).
- [0011] 본 발명의 프로세스의 희소 코딩 부분은 2개의 메인 컴포넌트(main component)들, 즉, (1) 딕셔너리 학습(dictionary learning) 및 (2) 딕셔너리 학습에 의해 정의된 좌표 공간 내의 수집된 데이터의 표현들을 컴퓨팅(computing)하고 그리고 딕셔너리의 벡터 공간(vector space) 내에 피팅(fit)되지 않는 데이터를 비정상성(anomaly)들로서 정의하는 것에 의한 비정상성 검출(anomaly detection)로 이루어진다. 본 발명에 따르면, 비정상성들(잔차들)은 정상 동작 범위를 초과하는 센서 관독들의 형태를 취한다. 본 발명의 바람직한 실시예에서, L-1 노름(L-1 norm)은 (종래의 L-2 노름 대신에) 잔차를 정의하기 위해 사용되는데, 그 이유는 L-1 노름("최소 절대 편차(least absolute deviation)" 또는 "최소 절대 에러(least absolute error)" 값으로 또한 알려져 있음)이 허위 포지티브 결과들을 산출할 가능성이 더 낮기 때문인데, 허위 포지티브 결과들은, 통상적으로 잡음있는 가스 터빈 센서 관독들을 연구할 때 문제가 된다.
- [0012] 희소 코딩에 의해 식별된 잔차들을 평가하기 위해 그리고 잔차들이 - 결함있는 센서가 존재하는 경우 - 결함있는 센서를 표시하는지를 결정하기 위해, 사후-프로세싱 기능(post-processing function)들이 사용된다. 일

예시적 사후-프로세싱 접근법은, 잔차들의 세트(set)의 모델(model)을 생성하기 위해 일-클래스(one-class) 지원 벡터 머신(SVM; Support Vector Machine)을 이용한다. 다른 사후-프로세싱 접근법은 잔차들의 그룹(group)의 특정 값들에 기반하여 임계값을 개발한다. 어느 경우이든, 시간-기반 판독들의 세트는 사용되고 있는 모델에 대해 비교되며, 시간에 따라 수집된 충분한 양의 데이터가 기준들을 충족하는 경우, "결함있는 센서" 검출이 선언된다. 그 후에, 수리 또는 교체되어야 할 특정 센서를 식별하는 "센서 고장" 메시지(message)가 가스 터빈 유지보수를 담당하는 직원에게 전송된다.

[0013] 일 실시예에서, 본 발명은 적어도 하나의 센서로부터 수집된 판독들에 기반하여 가스 터빈 센서 고장을 검출하기 위한 방법의 형태를 취하며, 방법은, (1) 정상 동작 상태들에 대해 알려진 센서 판독들과 연관된 값들을 정의하는 기저 벡터(basis vector)들의 디셔너리를 생성하는 단계; (2) L-1 노름 잔차 데이터(L-1 norm residual data)의 세트를 식별하기 위해, 생성된 디셔너리를 사용하여, 희소 코딩 프로세스를 센서 판독 데이터의 세트에 적용하는 단계; (3) 최대-값의 L-1 노름 잔차 데이터(largest-valued L-1 norm residual data)의 미리 결정된 서브세트(subset)를 비정상 센서 판독(abnormal sensor reading)들로 카테고리화(categorize)하기 위해 L-1 노름 잔차 데이터를 평가하는 단계; (4) 비정상 센서 판독들을 복수의 이전(prior) 센서 판독들과 비교하고 그리고 이전 센서 판독들 중 미리 정의된 개수의 이전 센서 판독들이 또한 비정상 센서 판독들로서 지정되는 경우, 지정된 비정상 센서 판독들을 센서 고장과 연관된 것으로 정의하는 단계; (5) 특정 센서를 식별하는 센서 고장 신호를 가스 터빈 직원(gas turbine personnel)에게 송신하는 단계; 및 (6) 고장난 센서를 서비스(service)로부터 제거하고 그리고 고장난 센서를 수리 또는 교체하는 단계를 포함한다.

[0014] 다른 실시예에서, 본 발명은 가스 터빈 센서의 고장을 검출하기 위한 시스템(system)으로서 설명되며, 시스템은 이러한 컴포넌트들, 즉, (1) 센서 판독들의 데이터베이스(database); (2) 센서 판독들의 데이터베이스와 통신하는 센서 모니터링 시스템 컴포넌트(sensor monitoring system component) - 센서 모니터링 시스템 컴포넌트는 프로그램 저장 디바이스(program storage device) 및 프로세서(processor)를 포함하고, 프로그램 저장 디바이스는 방법 단계들을 수행하기 위해 프로세서에 의해 실행가능한 프로그램 명령들의 세트를 고정형의 유형적 매체(fixed tangible medium)에 포함하고, 방법 단계들은, (a) 정상 동작 상태들에 대해 알려진 센서 판독들과 연관된 값들을 정의하는 기저 벡터들의 디셔너리를 생성하는 단계; (b) L-1 노름 잔차 데이터의 세트를 식별하기 위해, 생성된 디셔너리를 사용하여, 희소 코딩 프로세스를 센서 판독 데이터의 세트에 적용하는 단계; (c) 최대-값의 L-1 노름 잔차 데이터의 미리 결정된 서브세트를 비정상 센서 판독들로 카테고리화하기 위해 L-1 노름 잔차 데이터를 평가하는 단계; (d) 비정상 센서 판독들을 복수의 이전 센서 판독들과 비교하고 그리고 이전 센서 판독들 중 미리 정의된 개수의 이전 센서 판독들이 또한 비정상 센서 판독들로서 지정되는 경우, 지정된 비정상 센서 판독들을 센서 고장과 연관된 것으로 정의하는 단계 및 (e) 수리 또는 교체되도록 하기 위해, 고장난 센서를 식별하는 정보를 포함하는 센서 고장 신호를 가스 터빈 직원에게 송신하는 단계임 -; 및 (3) 비정상(abnormal)으로서 분류된 센서 판독들을 저장하기 위한 모니터링 시스템 데이터베이스를 포함한다.

[0015] 본 발명의 또 다른 실시예에서, 희소 코딩 프로세스에 의해 평가된 센서 판독들은 (고장이 이미 발생한 상태와는 대조적으로) 임박한 센서 고장을 예측하기 위해 사용될 수 있다.

[0016] 본 발명의 다른 양상들 및 실시예들 그리고 추가의 양상들 및 실시예들은 다음의 논의의 과정 동안 그리고 첨부 도면들을 참조하여 명백해질 것이다.

도면의 간단한 설명

[0017] 이제 도면들을 참조하면, 여러 도면들에서 동일한 숫자들은 동일한 부분들을 표현한다:

도 1은 결함 센서들을 검출하는 본 발명의 시스템 및 방법이 사용될 수 있는 예시적 가스 터빈 환경의 간략화된 도면이고;

도 2는 수집된 센서 판독들을 평가하기 위해 그리고 하나 또는 그 초과 의 센서들이 결함있는지 및 수리 또는 교체될 필요가 있는지를 결정하기 위해 활용될 수 있는 본 발명의 예시적 프로세스의 흐름도이고;

도 3은, 센서 고장들을 검출하기 위해 L-2 노름(L-2 norm)을 사용하는 것과 연관된 문제들을 예시하는, L-2 노름을 희소 코딩 프로세스에 적용함으로써 컴퓨팅된(computed) 잔차들의 다이어그램(diagram)이고;

도 4는 L-1 노름을 희소 코딩 프로세스에 적용함으로써 본 발명의 교시들에 따라 컴퓨팅된 잔차들의 다이어그램이고;

도 5는, 어느 잔차들이 적절하게 "비정상"인지를 결정하고 센서의 고장을 표시하기 위해, 일-클래스 지원 벡터

머신(SVM; Support Vector Machine) 모델 접근법을 사용하여 L-1 노름 잔차들을 평가하기 위한 방법의 흐름도이고; 그리고

도 6은 L-1 노름 잔차들을 평가하고 센서 고장의 존재를 검출하기 위한 대안적 방법의 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0018] 도 1은, 플랜트 전기 부하(plant electric load)(3)를 공급하는 발생기(2)를 갖는 통상의 가스 터빈 전력 플랜트(1)의 간략화된 도면이다. 발생기(2)는 가스 터빈 엔진(gas turbine engine)(5)에 의해 동력이 공급되는 축(4)에 의해 구동된다. 가스 터빈 엔진(5) 그 자체는, 압축기(5.1), 연소 섹션(combustion section)(5.2), 터빈(5.3) 및 아마도 조정가능한 입구 베인(inlet vane)들(5.4)의 세트를 포함한 많은 수의 개별 컴포넌트들로 이루어진다.
- [0019] 연료는 밸브(valve)(6)를 통해 연소 섹션(5.2)에 공급된다. 가스 터빈 전력 플랜트(1)의 수용가능한 동작을 유지하기 위해, 다수의 센서들(7)이 사용되어 다양한 컴포넌트들의 동작을 모니터링(monitor)하고, 측정된 센서 판독들을 개별 제어 모듈(control module)(8)에 전달한다. 제어 모듈(8)은 가스 터빈 전력 플랜트(1)와 병치될(co-located) 수 있거나 또는 터빈 그 자체로부터 떨어져 있을 수 있다. 도 1의 다이어그램에서, 센서들(7)은 연소기 유입 공기 센서(combustor inlet air sensor)(7.1), 연소 온도 센서(7.2), 및 블레이드 경로 온도 센서(blade path temperature sensor)(7.3)를 포함한다. 가스 터빈의 성능을 모니터링하기 위해 사용되는 더 많은 센서들이 존재하여서, 온도, 압력, 회전, 진동 등과 같은 상태를 측정한다는 것이 이해되어야 한다. 실제로, 주어진 가스 터빈 전력 플랜트에 따라 200개에 가까운 상이한 센서들이 활용될 수 있는 것이 가능하다.
- [0020] 제어 모듈(8)은 센서들(7)로부터 입력을 수신하고, 제어 신호들을 당해 기술 분야에 알려진 바와 같은 밸브들, 모터(motor)들, 및 액추에이터(actuator)들에 송신한다. 제어기는, 당해 기술 분야에 알려진 바와 같은 연관된 하드웨어(hardware) 및 소프트웨어(software)를 갖는 하나 또는 그 초과 프로세서들을 하나 또는 그 초과 위치들에 포함할 수 있다.
- [0021] 가스 터빈의 성능이 정확한 센서 판독들에 종속되는 것을 고려하면, 센서가 고장나기 시작하여 부정확한 판독들을 전달할 때를 인지할 수 있는 것이 중요하다. 위에서 언급된 바와 같이, 센서 거동을 이해하는 중요한 양상은, 일회성의 결함있는 판독(일회성의 결함있는 판독은 다양한 환경 팩터들로 인한 것일 수 있고 따라서 무시될 수 있음)과 일련의 계속되는 결함있는 판독들(일련의 계속되는 결함있는 판독들은 프리젠테이션(presentation)에서 더 미묘할 수 있음) 간을 판별하는 능력을 포함하는 것인데, 그 이유는 일련의 계속되는 결함있는 판독들의 상태가, 고장나서 교체될 필요가 있는 센서로부터 초래될 가능성이 더 많기 때문이다. 더욱이, 위에서 언급된 바와 같이, 가스 터빈 센서들에 의해 수집된 통상의 아날로그 센서 판독(analog sensor reading)들은 상대적으로 큰 잡음 컴포넌트를 포함하는 경향이 있어서, "잡음있는 데이터"와 "결함있는 센서" 간을 구별하는 능력을 제한한다.
- [0022] 기본 임계치 센서 판독을 "알람(alarm)"으로서 사용하는 다양한 종래 기술의 기법들과 대조적으로, 본 발명의 방법론은 센서 판독들의 시계열 데이터 세트(time series data set)에 대한 분석을 수행한다. 수용가능한 센서 판독들의 유니버스(universe)를 적절하게 특징짓는 디서너리를 유도하기 위해 최소 코딩 기법이 사용되며, 여기서 기저 벡터들의 정의된 경계들 외측에 있는 임의의 판독들("잔차(residual)"로 지칭됨)은 잠재적 문제로 분류된다. 상황을 "결함있는 센서"로서 정의하기 위한 시간의 함수로써 충분한 양의 비정상 판독들이 존재하는지를 결정하기 위해, 잔차들의 모델이 생성되어 사용된다.
- [0023] 그러므로, 본 발명에 따르면, 제어 모듈(8)의 성능들은 센서 모니터링 시스템(10)을 포함하도록 확장되며, 센서 모니터링 시스템(10)은 아래에서 상세하게 논의되는 바와 같이, 수집된 센서 판독들을 분석하고 센서 판독들에서 특정 센서의 고장의 표시자인 임의의 텔-테일 변화(tell-tale change)들을 찾기 위해 활용될 수 있거나 또는 다음 수 일 내에 센서 고장이 발생할 가능성이 있다는 것을 함리적으로 예측할 수 있다. 아래에서 상세하게 논의될 바와 같이, 센서 모니터링 시스템(10)은 (수년 전으로 되돌아갈 수 있는) 모든 이력 센서 판독들의 데이터 베이스(12)를 포함하는데; 실제로 이 데이터베이스는 종래의 제어 모듈(8)의 부분을 형성할 수 있다. 센서 모니터링 시스템(10)은 메모리(memory)(16) 및 프로세서(18)를 갖는 프로세서 컴포넌트(14)를 포함하며, 이들 엘리먼트(element)들은 센서 판독 데이터의 실제 평가를 수행하고 센서들 그 자체의 성능을 모니터링하는 데 사용된다. 그 후에, "비정상" 센서 판독의 경보의 형태를 취하는 센서 모니터링 시스템(10)으로부터의 출력은, 센서가 교체 또는 수리될 필요가 있는지의 여부를 결정하는 데 있어서 플랜트 직원에 의해 사용된다. 도 1에 도시된 바와 같은 센서 모니터링 시스템(10)의 특정 아키텍처(architecture)는 단지 예시적인 것으로 고려되며,

아래에서 상세하게 설명되는 바와 같은 본 발명의 방법을 수행하도록 다양한 다른 구성들의 하드웨어 및 소프트웨어 컴포넌트들이 구성될 수 있다.

- [0024] 도 2는 수집된 센서 판독들을 평가하고 하나 또는 그 초과 센서들이 결합있는지 및 수리 또는 교체될 필요가 있는지를 결정하기 위해 센서 모니터링 시스템(10)에 의해 활용될 수 있는 본 발명의 예시적 프로세스의 흐름도이다.
- [0025] 프로세스는, 단계(200)에 도시된 바와 같이, 실제 센서 판독 데이터를 수집하고 데이터를, 데이터가 나중에 리트리빙되고(retrieved) 분석될 수 있는 엘리먼트에 저장함으로써 시작된다. 예컨대, 도 1의 제어기(8)는 이러한 기능들을 수행하기 위해 메모리 엘리먼트들 및 프로세싱 성능들을 포함하는 것으로 가정된다. 일단 데이터가 수집되면, 예비 단계는, 제거되지 않으면 분석의 결과들을 왜곡(skew)할 명백한 이상점(outlier)들 및 리던던트(redundant) 데이터를 제거하기 위해 센서 판독들을 "사전프로세싱(preprocessing)"하는 단계(단계(210))를 수반할 수 있다. 이 사전프로세싱 단계가 의무적인 것이라고 고려되지 않지만(그리고 흐름도에서 점선 박스(box)로서 도시됨), 본 발명의 바람직한 실시예에서, 사전프로세싱 단계는 효율적인 방식으로 정확한 해결책(solution)에 도달하기 위해 활용된다.
- [0026] 단계(220)는 수집된 데이터 세트에 대해 "최소 코딩"을 수행하는 실제 프로세스와 연관된다. 아래에서 더 상세하게 논의될 바와 같이, 최소 코딩은, 데이터 세트를 정확하게 표현하기 위해(이 경우, 센서 판독들을 정확하게 표현하기 위해) 사용될 수 있는 기저 벡터들의 세트를 정의하는 것을 포함한다. "잔차들"로 지칭되는, 최소 코딩 벡터 공간 내에 피팅(fit)되지 않은 소수의 데이터 포인트(few data point)들은, 잡음이 있을 수 있거나 또는 그렇지 않으면 결합있는 센서들과 연관될 수 있는 그러한 데이터이다. 데이터가 벡터 공간 외측에 있는 레벨(level)을 결정하기 위해 다양한 상이한 벡터 노름들이 사용될 수 있지만, 센서 판독들에 기인하는 잡음있는 데이터는 L-1 노름(즉, 최소 절대 편차 노름(least absolute deviation norm))을 사용하는 것에 제일 적합하다.
- [0027] 그 후에, 단계(220)에서 식별된 바와 같은 잔차들은, 데이터 포인트들이 단지 잡음있는 데이터인지 또는 정말로 결합있는 센서와 연관되었는지를 궁극적으로 결정하기 위해 추가로 평가된다. 다음 단계(도 2의 흐름도에서 단계(230)로 도시됨)는 잔차들을 2개의 그룹들, 즉, 트레이닝 세트(training set)로서 사용되는 제 1 그룹 및 테스트 세트(test set)로서 사용되는 제 2 그룹으로 분할한다. 그 후에, 트레이닝 잔차들은 아래에서 논의될 바와 같이, 잔차들의 그룹을 평가하고 그리고 소수의 가장 큰 잔차들을 센서의 "비정상" 동작과 연관된 것으로 정의하기 위한 한계치를 결정하는 데 적합한 모델을 개발하기 위해 사용된다(단계(240)).
- [0028] 일단 모델의 파라미터(parameter)들이 식별되면, 테스트 데이터에 대해 최소 코드 프로세싱(sparse code processing)을 수행함으로써 생성된 잔차들이 이 모델을 통과하며(단계(250)), 여기서 비교가 수행되고, 테스트 되고(tested) 있는 현재의 "잔차"가 모델 내에 있는지 또는 모델 외측에 있는지의 여부의 결정이 이루어진다(단계(260)). 그 후에, 모델 외측에 있는 그러한 잔차들은 "비정상"으로 식별되고, 그와 같이 라벨링되고(labeled), 추가의 사용을 위해 메모리에 저장된다.
- [0029] 프로세스의 이 시점에서 분석의 시간 컴포넌트가 도입된다. "비정상" 데이터 포인트의 디스커버리(discovery)에 따라, 프로세스의 다음 단계(단계(270))는 다수의 이전의 시간 포인트들에 걸쳐 데이터 포인트들을 조사(review)하고 이들 데이터 포인트들 중 상당한 수가 또한 "비정상"으로서 식별되었는지를 확인하는 것이다. 또한, 이 시점에서, 주어진 시간 기간에 걸쳐 발견되도록 요구되는 "비정상" 판독들의 수를 평가하는 데 있어서 상이한 측정들이 사용될 수 있다.
- [0030] 시간의 함수로써 센서 판독 데이터의 상당한 퍼센티지(percentage)가 계속해서 "비정상"으로 플래그되는(flagged) 경우(단계(280)), 연관된 센서가 결합이 있을 수 있고 실제 디바이스를 평가할 필요가 있다는 결정이 이루어진다(단계(290)). 그렇지 않으면, 현재의 "비정상" 판독은 메모리 엘리먼트에 유지되고, 센서 판독들을 모니터링하는 프로세스가 계속된다(단계(300)). 센서가 고장나기 시작함에 따라 센서 판독 데이터가 저하되기 시작하여 더 많은 수의 잔차들을 생성할 수 있기 때문에, 도 2의 흐름도에서 간략히 서술된 바와 같은 프로세스는 또한, 센서 고장들의 발생 전에 센서 고장들을 예측하는 데 유용하다는 것이 이해되어야 한다.
- [0031] 다음의 논의는, 특히, 최소 코딩 기법이 본 발명의 가스 터빈/센서 환경에 적용됨에 따른 최소 코딩 기법의 간략한 개요를 포함하며, 그 후에 잔차들을 분석하고 시간 팩터를 분석에 도입하기 위한 모델을 개발하는 데 사용될 수 있는 2개의 상이한 사후-프로세싱 기법들의 논의가 뒤따른다. 그 후에, 실제 가스 터빈 센서 판독들에 대한 본 발명의 방법론의 적용의 예가 뒤따른다.

[0032] "희소 코딩" 프로세스는 일반적으로, 입력 신호의 간결한 표현을 발견하기 위한 알고리즘(algorithm)들의 클래스(class)와 관련된다(이 경우, 입력 신호는 가스 터빈 센서 판독들의 형태를 취함). 특히, 신호 분석 기법으로서의 희소 코딩은, 효율적인 방식으로 데이터를 표현하는 오버-컴플리트 베이스(over-complete base)들의 세트들(때때로 기저 벡터들의 세트로 또한 지칭됨)을 학습하기 위한 방법론의 부분이다. "희소(sparsity)"는, (소수의 큰 계수들을 제외한) 변환 계수들의 대부분이 제로(zero)로 근사화될 때 존재하는 제한된 수의 비-제로 컴포넌트들을 나타낸다. 본 발명에 따르면, 연구되고 있는 데이터는 시계열 센서 판독들(각각의 센서에 대해 수집된 데이터의 상이한 세트)이다. 가스 터빈 센서 데이터를 연구하기 위한 종래의 기법들(이를테면, 위에서 언급된 PCA 접근법)과는 대조적으로, 희소 코딩은 기저 벡터들이 입력 데이터에서 구조들 및 패턴(pattern)들(뿐만 아니라 비정상성들)을 더 양호하게 캡처(capture)하는 능력을 제공한다(희소 코딩이 기저 벡터들의 오버-컴플리트 세트이기 때문임). 희소 코딩은 주어진 입력 데이터 세트 X 가 디셔너리 $D = [d_1, d_2, \dots, d_i] \in \mathbb{R}^{m \times k}$ 에서 하나 또는 그 초과 기저 벡터들의 선형 조합으로서 표현되도록 허용한다.

[0033] 본 발명에 따르면, 결합있는 센서 판독들은 "비정상성들"로서 정의되며, 비정상성 검출과 함께 희소 코딩 프로세스를 활용함으로써 컴퓨팅될(computed) 수 있다. 희소 코딩은 2개의 메인 컴포넌트들, 즉, 디셔너리 학습 컴포넌트(dictionary learning component) 및 "디셔너리 공간(dictionary space)"을 형성하는 기저 벡터들로 입력 데이터를 표현하기 위한 좌표 컴포넌트(coordinate component)로 이루어진다. 디셔너리 학습 부분의 경우, 아래의 최적화 문제가 해결된다:

[0034]
$$\min_{D, X} \|Y - DX\|_F^2 + \alpha \|X\|_1, s.t. \|D\| \ll 1 \quad (1)$$

[0035] X 의 작은 노름에 도달하기 위해 디셔너리 D 가 임의적으로 크게 성장하지 않는다는 것을 보장하기 위하여 제약 $\|D\| \ll 1$ 이 부과된다. 디셔너리 학습 문제의 제 1 항은 "재구성" 항(즉, 기저 벡터들의 세트에 데이터를 표현함)이고, 제 2 항은 (α 의 선택된 값의 함수로써) 표현에서 희소성(sparseness)을 촉진하기 위해 포함된다.

[0036] 최적화 문제는 D 와 X 에서 양면볼록(biconvex)인데; 즉, 하나의 변수가 고정될 때, 나머지 최적화 문제는 볼록 최적화 문제이다. 그러므로, 최적의 해결책으로의 수렴이 보장되도록 그리고 승수들의 교번 방향 방법(ADMM; Alternating Direction Method of Multipliers)과 같은 접근법들(또는 상대적으로 또한 강건하고 데이터의 임의적-스케일(arbitrary-scale)의 최적화를 제공할 수 있는 다른 접근법들)이 바람직한 것으로 간주되도록, 각각의 변수에 대해 교번적인 방식으로 해결함으로써 수용가능한 결과가 제공되는 것이 뒤따른다. 궁극적으로, 센서 판독 데이터를 기저 벡터들의 선형 조합으로서 정의하기 위해 사용될 수 있는 기저 벡터들의 디셔너리 D 가 획득된다. 실제 데이터에 기반하여 벡터들이 실시간으로 구축되는(built) 온라인-학습 디셔너리(online-learning dictionary)를 포함한 임의의 적절한 디셔너리 학습 알고리즘이 활용될 수 있다. 부가적으로, 디셔너리를 초기화하기 위해 임의적인 값들을 사용하기보다는, 바람직한 접근법은 k -평균 알고리즘(k -means algorithm)을 사용하여 데이터 클러스터(data cluster)들의 적절한 센터(center)들을 발견하여, 그 센터들을 프로세스를 초기화하는 데 사용하는 것이다. 이 기법은 학습 프로세스에 대한 수렴의 속도를 개선하는 것으로 고려된다.

[0037] 이 디셔너리는, "정상 값"의 임의의 착신 판독(incoming reading)이 기저 벡터들의 선형 조합으로서 정확하게 표현될 수 있는 수용가능한 센서 판독들의 유니버스(universe)로 생각될 수 있다. 본 발명에 따르면, 이 디셔너리는, 연관된 센서의 성능을 모니터링하고 디셔너리 공간 외측에 있는 센서 판독들(이들 판독들은 "잔차들"로 지칭되며, 하기의 논의에서 E 로 표기됨)의 존재를 검출하기 위해 희소 코딩 프로세스와 함께 사용된다. 즉, 주어진 디셔너리 D 에 대해, 본 발명의 프로세스는 아래의 관계식을 사용하여 착신 센서 판독들을 디셔너리 D 에 대해 평가함으로써 진행된다:

[0038]
$$\min_{E, X} \|Y - DX - E\|_F^2 + \alpha \|X\|_{1,2} + \beta \|E\|_{1,1} \quad (2)$$

[0039] 여기서 $\|X\|_{1,2}$ 는 행렬 X 의 각각의 행의 L_2 노름의 합이고, $\|E\|_{1,1}$ 는 행렬 X 의 각각의 열의 L_1 노름 잔차들의 합이며, α 및 β 는 작은 포지티브 스칼라(small positive scalar)들이다. 이 접근법은 또한 "다중 측정 벡터(multiple measurement vector)"(MMV) 접근법으로 알려져 있다. 항 $\beta \|E\|_{1,1}$ 은 "잔차"이며, 착신

센서 데이터가 수용 값들 너머로 벗어나서 센서가 아마도 "결함있음"으로서 평가될 필요가 있는지를 결정하는데 사용된다. β 팩터는, 선택된 센서 판독들을 잔차로서 식별하는 감도의 레벨을 조정하기 위해 모니터링 직원에 의해 사용될 수 있다.

[0040] 도시된 바와 같이, 비정상성 검출을 위해 사용되는 종래의 L-2 노름보다는, L-1 노름이 이러한 "잔차"를 정의하는데 사용된다. 이는, L-2 노름이 상당한 수의 큰 잔차들을 (그리고 결과적으로, 결함있는 센서들의 관점들에서 상당한 수의 허위 포지티브들을) 생성하기 때문이다. 아래에서 정의되는 바와 같은 데이터포인트 값(datapoint value)들의 합인 L-1 노름은, 큰 잡음 팩터를 포함할 수 있는 데이터를 평가할 때 바람직한 대책인 것으로 간주되며, 목표는 잡음 예러와 비정상 센서 판독들 간의 차이를 판별할 수 있는 것이다. 논의된 바와 같은 L-1 노름의 사용은 잔차에서 희소성을 촉진할 뿐만 아니라 데이터에 대한 더 양호한 피팅(fit)을 제공한다. 그에 반해, 잡음있는 데이터(이 경우, 여기서는 가스 터빈 센서 데이터)를 처리할 때 L-2 노름이 사용될 경우, 결과적인 잔차들은 (선호되는 더 작은 잔차 값들 대신에) 매우 많으며 상대적으로 큰 경향이 있을 것이다.

[0041] 도 3 및 도 4는 잔차들을 생성하는데 있어서 L-2 노름을 사용하는 것(도 3에 도시되는 바와 같음)과 L-1 노름을 사용하는 것(도 4에 도시되는 바와 같음) 간의 결과들의 차이를 예시한다. L-2 노름이 상당히 많은 수의 데이터 포인트들을 "잔차들"로서 식별하는 것을 초래하고, 이는 바람직하지 않은 수의 허위 포지티브 "결함있는 센서" 알람들을 초래할 수 있다는 것이 명백하다. 반면에, L-1 노름을 사용함으로써 생성된 잔차들은 더 적고 훨씬 더 뚜렷하다. 이들 그래프(graph)들과 관련된 "실형" 설명 세부사항들은 아래에서 상세하게 설명될 것이다.

[0042] 본 발명의 방법론의 논의를 다시 참조하면, 수학적(2)으로 표현되는 바와 같은 문제는 ADMM 최적화 기법을 적용함으로써 해결된다. 이 경우, 잔차들 E^k 의 세트의 형태를 취하는 해결책은 노름으로부터의 경계들 외측에 있는 센서 판독들의 세트이며, 센서의 결함있는 동작과 연관될 수 있다. 아래의 목록에서 보여지는 바와 같이, 반복적 방식(아래의 단계 3)으로 잔차 E가 업데이트될(updated) 때 종래의 수축 연산자(shrinkage operator)(아래의 단계 2에서 사용되는 바와 같음)가 엘리먼트-방식 "소프트 임계처리(soft thresholding)" 연산자로 대체되도록, MMV 알고리즘(위에서 양면블록 변수들 D 및 X를 논의할 때 논의된 바와 같음)이 변경된다. 소프트 임계 처리는, 여전히 데이터를 가능한 한 깨끗하게 분할하지만, 최대 마진(maximum margin)을 수정하고 그리고 일부 잘못 라벨링된(mislabeled) 데이터가 생성되게 하는 "느슨한(slack)" 변수를 도입하는 수정된 절차이다. 아래의 목록은 최소 코딩의 L-1 잔차들을 컴퓨팅(compute)하기 위해 사용되는 알고리즘을 상세하게 보여주며, 여기서 바로 아래에 포함된 테이블(table)은 알고리즘에 포함된 파라미터들을 정의한다:

파라미터	정의
Z	보조 변수
E	잔차
M	라그랑주 승수
μ	페널티 계수 (ADMM 공식에서 동등 제약과 연관된 페널티를 제어)
ρ ($\rho > 1$)	μ 를 업데이트하기 위한 ADMM 알고리즘의 파라미터
ϵ	알고리즘을 종결시키기 위해 Z 내지 X의 근사를 측정하는데 사용되는 에러 허용한계

[0043]

[0044] 최소 코딩의 L-1 잔차들을 컴퓨팅하기 위한 알고리즘 (잔차들은 가능하게는 잠재적인 센서 고장을 표시하는 것으로 분석됨)

[0045] 해결:
$$\min_{E,X} \|Y - DX - E\|_F^2 + \alpha \|X\|_{1,2} + \beta \|E\|_{1,1}$$

[0046] 입력: 신호들 $X \in \mathfrak{R}^{N \times L}$, 디크서너리 $D \in \mathfrak{R}^{N \times M}$

[0047] 초기화: $k=0$; $Z^0, E^0, M^0, \mu^0 > 1, \epsilon$ 로 설정

[0048] 수렴까지 반복:

[0049] 1.
$$X^{k+1} = \arg \min_X L(X, Z^k, E^k, M^k, \mu^k) = (D^T D + \mu^k I)^{-1} (D^T (Y - E^k) + M^k + \mu^k Z^k)$$

[0050] 여기서 L은 증강된-라그랑지안(Augmented-Lagrangian)이고 아래와 같이 정의됨:

[0051]
$$L(X, Z, E, M, \mu) = \frac{1}{2} \|Y - DX - E\|_F^2 + \alpha \|Z\|_{1,2} + \beta \|Z\|_{1,1} + \langle M, Z - X \rangle + \frac{\mu}{2} \|Z - X\|_F^2$$

[0052] 2.
$$Z^{k+1} = \arg \min_Z L(X^{k+1}, Z, E^k, M^k, \mu^k) = \arg \min_Z \frac{1}{2} \|P - Z\|_F^2 + \gamma \|Z\|_{1,2}$$

[0053] 여기서 $P = X^{k+1} - \frac{1}{\mu^k}$ 임. 이 최적화 문제는 행 수축 연산자(row shrinkage operator)를 사용함으로써 해결될 수 있음:

[0054]
$$Z^{k+1}(j, :) = \begin{cases} \frac{\|P(j, :)\|_2 - \gamma}{\|P(j, :)\|_2} P(j, :) & \gamma < \|P(j, :)\|_2 \text{ 인 경우} \\ 0 & \text{그렇지 않은 경우} \end{cases}$$

[0055] 3.
$$E^{k+1} = \arg \min_E L(X^{k+1}, Z^{k+1}, E, M^k, \mu^k) = \arg \min_Z \frac{1}{2} \|Q - E\|_F^2 + \beta \|E\|_{1,1}$$

[0056] 여기서 $Q = Y - DX^{k+1}$ 임. 이 최적화 문제는 엘리먼트-방식 소프트 임계처리 연산자 $E^{k+1} = S_\beta(Q)$ 를 사용함으로써 해결될 수 있으며, 여기서 아래와 같음.

[0057]
$$S_\beta(a) = \begin{cases} a - \beta & a > \beta \text{ 인 경우} \\ 0 & |a| \leq \beta \text{ 인 경우} \\ a + \beta & a < -\beta \text{ 인 경우} \end{cases}$$

[0058] 4.
$$M^{k+1} = M^k + \mu^k (Z^{k+1} - X^{k+1})$$

[0059] 5.
$$\mu^{k+1} = \rho \mu^k$$

[0060] 6.
$$k = k + 1$$

[0061] 7.
$$\frac{\|Z^k - X^k\|_F^2}{\|X^k\|_F^2} < \epsilon$$
 인 경우, 중지됨.

[0062] 출력: E^k

- [0063] 이 알고리즘에 의해 생성되고, "잔차들"로서 정의된 E^k 의 값들은 가스 터빈 센서의 결합있는 동작과 연관된 잠재적 데이터 포인트들이다. 도 2의 흐름도를 다시 참조하면, 생성된 잔차들은 단계(220)로부터 출력된다. 이들 잔차들을 평가하고 그리고 시간 컴포넌트를 고려하기 위해, 본 발명에 따라 활용될 수 있는 적어도 2개의 상이한 접근법들이 존재한다.
- [0064] 제 1 접근법에서, 방사형 기저 함수(RBF; radial basis function) 커널(kernel)을 사용하여 일-클래스 지원 벡터 머신(SVM; support vector machine) 모델을 학습하기 위해 트레이닝 잔차가 사용된다. 일반적으로 말하면, SVM 모델은 데이터(이 경우, 잔차들)를 분석하고 패턴들을 인식하기 위해 학습 알고리즘들과 관련하여 사용된다. SVM은 개별 카테고리(category)들의 모델을 구축하고 그 후에 새로-도착하는 데이터를 적절한 카테고리에 할당한다. 여기서 사용되는 바와 같은 "일-클래스" SVM은 단지 단일 카테고리만을 생성하고, 새로-도착하는 데이터를 평가하도록 기능하고, 데이터가 카테고리 내에 있는지의 여부를 결정한다. SVM 알고리즘의 효율성을 증가시키기 위해, 전체 세트 대신에 트레이닝 잔차들의 샘플링(sampling)을 사용하는 것이 가능하다. 그 후에, SVM "일 클래스" 모델은 새로-도착하는 잔차들에 적용된다. 이 모델에서, 잔차 데이터 대부분은 생성된 모델의 경계들 내에 속할 것이고, 상대적으로 소수의 값들만이 "이상점들"로 남겨진다. 그 후에, 이들 이상점들은, 센서 그 자체의 추가의 평가를 프롬프트(prompt)할 수 있는 "비정상" 센서 판독들로서 정의된다. 실제로, 비정상 판독들은, 센서 고장이 이미 발생했다고 표시할 수 있거나, 또는 대안적으로, 센서가 고장나기 시작하고 주의가 필요하다는 예측으로서 사용될 수 있다.
- [0065] 허위 포지티브들의 수를 감소시키기 위해, 이전의 T 시간 단계들에 대한 센서 판독 데이터를 조사하는 것이 제안된다. 조사되는 데이터의 높은 퍼센티지(p%로 표기됨)가 또한 (생성된 일-클래스 SVM 모델에 따라) 비정상으로서 정의되는 경우, 현재의 데이터 포인트는 개연성있는 "결합있는 센서"를 경보하는 것으로 정의된다.
- [0066] 도 5의 흐름도는 센서 결합 검출의 이러한 예시적인 프로세스를 간략히 서술한다. 흐름도는, 특정 센서에 대한 데이터가 조사되고 있고 최대 센서 판독 값의 20% 미만인 판독들 및 이상점들을 제거하기 위해 데이터가 클리닝된다고(cleaned) 가정한다.
- [0067] 프로세스는 단계(500)에서, 다양한 가스 터빈 센서들로부터 원시 데이터(raw data)를 수집함으로써 시작된다(위에서 논의된 바와 같이, 이 데이터는 제어기(8)의 메모리 컴포넌트 내에 저장됨). 논의를 위해, 터빈의 생성된 출력 전력(MW)과 연관된 센서가 평가되는 것으로 가정된다. 단계(510)에서 도시되는 바와 같이, 데이터는 바람직하게, 이 파라미터와 연관되지 않은 센서들뿐만 아니라, 데이터가 전력 센서 출력과 강하게 상관되지 않은 임의의 센서들을 제거하기 위해 사전-프로세싱된다(pre-processed).
- [0068] 일단 사전프로세싱되면(preprocessed), 나머지 데이터는 위에서 간략히 서술된 바와 같이 희소 코딩 프로세스를 겪는다. 도 5에 도시된 바와 같이, 이 프로세스는 먼저 이 데이터에 대한 기저 벡터들의 디셔너리를 구축함으로써(단계(520)), 그리고 그 후에 수확식(2)에 도시된 바와 같은 비정상성 검출과 함께 희소 코딩 알고리즘을 센서 판독 데이터에 적용하여, 이 데이터로부터 L-1 노름 잔차들을 컴퓨팅함으로써(단계(530)) 시작된다. 디셔너리를 구축하는 다양한 방법들이 단계(520)에서 사용될 수 있지만, 효율적인 방식으로 수행되는 바람직한 프로세스는, 데이터를 클러스터링(clustering)하는 단계(단계(522)), 클러스터들의 센터들을 발견하기 위해 k-평균 알고리즘을 사용하는 단계(단계(524)), 및 이들 "센터" 값들을 이용하여 디셔너리를 초기화하는 단계(526)를 포함한다. 일단 구축되면, 디셔너리는 L-1 노름 잔차들을 식별하기 위해 단계(530)의 희소 코딩 프로세스에서 사용된다. 일단 식별되면, L-1 노름 잔차들은 그 후에, 이들 잔차들 중 어느 것이, 결합있는 센서를 표시하는 정말로 "비정상" 센서 판독들인지를 결정하기 위해 데이터의 시간-기반 조사를 포함하여 추가로 평가된다.
- [0069] 도 5를 참조하면, L-1 노름 잔차들의 세트는 먼저, 분석의 나머지에서 수반되는 데이터포인트들의 수를 감소시키기 위해 샘플링될(sampled) 수 있다(단계(540)). 이 단계는 선택적이지만, 다음의 단계(550)에서 사용되는 SVM 알고리즘의 효율성을 개선하는 것으로 고려된다. 실제로, SVM 알고리즘은 그 후에, 잔차들의 샘플링된 그룹을 "트레이닝 데이터"로서 사용하고, 적절한 일-클래스 모델을 구축하며, 적절한 일-클래스 모델은 그 후에, 새로-도착한 잔차들을 클래스 내에 있는 것으로 또는 클래스 외측에 존재하는 것(이상점 데이터는 본 목적들을 위해 "비정상"으로서 정의됨)으로 분류하는 데 사용된다.
- [0070] 개발된 일-클래스 모델의 정확도를 검증하기 위해(그리고 도 2의 일반적 흐름도에서 간략히 서술된 바와 같이), 테스트 데이터의 세트는 일-클래스 SVM 모델을 통과한다. 도 5를 참조하면, 테스트 데이터는 (단계(520)에서 구축된 디셔너리를 사용하여) L-1 잔차들의 테스트 데이터의 세트를 컴퓨팅하기 위해 프로세싱된다(단계(600)). 그 후에, L-1 잔차들의 테스트 데이터 세트는 단계(550)에서 생성된 SVM 모델을 통과한다(단계(610)로서

도시됨). (즉, 단계(620)에서의 평가 동안에 SVM 모델의 경계들 외측에 있는) "비정상"으로서 식별되는 임의의 데이터 포인트들이 수집 및 저장된다(단계(630)).

[0071] 본 발명에 따르면, 분석과 연관된 시간 팩터가 이제 도입되며, 여기서 단계(640)에서 이전의 T 시간 단계들의 세트에 대한 데이터의 조사가 수행된다("T"는 분석을 수행하는 개인에 의해 결정된 값임). 그 후에, 단계(650)에서 결정이 이루어지는데, 여기서 이전의 T 시간 단계들 중 p% 초과가 또한 "비정상" 데이터 포인트들을 포함한 경우(또한, p는 사용자에 의해 선택된 값임), 연관된 센서는 아마도 "결함있음"으로 정의되며, 실제 센서의 평가가 적절하다(단계(660)). 그렇지 않고, 이전의 T 시간 단계들 중 p% 미만이 이들 비정상 판독들을 갖지 않는 경우, 현재의 비정상 판독은 잡음있는 판독, 가짜 에러 등에 기인할 수 있다고 결론내려질 수 있으며; 어느 경우이든, p% 임계치가 초과되지 않으면, 센서가 적절하게 동작하고 있고 어떠한 추가의 평가도 요구되지 않는 것으로 추정될 수 있다(단계(670)).

[0072] 일-클래스 SVM 모델을 사용하는 대신에, 잔차 데이터를 추가로 연구하고 데이터 중 임의의 데이터가 관련 센서의 결함있는 성능과 연관될 수 있는지를 확인하기 위해, (잔차들을 지정하기 위해 L-1 노름을 활용하는) 최소 코딩 방법론을 다른 종류들의 분석들과 조합하는 것이 가능하다. 도 6은 하나의 이러한 대안적 분석 프로세스의 흐름도이며, 이러한 경우 여기서, 착신 센서 판독들이 결함있는 센서를 표시하는지를 결정하기 위해 일-클래스 SVM 모델 대신에, 임계치 분석이 사용된다. 도 6의 흐름도에서 도시되는 바와 같이, 프로세스의 초기 단계들은 도 5에 도시된 것들과 동일한데; 즉, 트레이닝 데이터를 사전프로세싱하고(단계(510)), 기저 벡터들의 디셔너리를 구축하고(단계(520)), 최소 코드 프로세싱을 활용하여(단계(530)), L-1 잔차들의 세트를 컴퓨팅한다.

[0073] 일단 트레이닝 데이터와 연관된 잔차들이 컴퓨팅되면, 도 6에 도시된 바와 같은 프로세스는, 단계(700)에서, 트레이닝 데이터에서 식별된 잔차들의 세트를 분석하고, 임계치를 초과하는 임의의 값이 "비정상" 데이터 포인트로서 지정되도록 그룹에 대한 적절한 임계값을 결정함으로써 계속된다. 이 임계치는 모든 센서들에 대해 균일할 수 있거나 또는 센서 단위 기반으로 정의될 수 있다. 예컨대, 5%의 가장 큰 잔차 값들이 "비정상"으로서 정의되고 나머지 잔차들은 무시되도록, 95%의 임계값이 사용될 수 있다.

[0074] 프로세스의 이 시점에서, 단계(520)에서 구축된 것과 동일한 디셔너리를 사용하여 그리고 단계(530)에 의해 예시된 동일한 최소 코딩 프로세스를 수행하여, 테스트 데이터(testing data)에 대한 L-1 잔차들이 다음에 결정된다(단계(710)). 일단 테스트 데이터에 대한 L-1 잔차들이 컴퓨팅되면, 이들은 단계(700)에서 정의된 임계치에 대해 (단계(720)에서) 비교된다. 주어진 테스트 데이터 잔차가 (결정(725)에서) 임계치 미만인 경우, 데이터는 단계(730)에서 "잡음있는" 데이터(또는 어떠한 추가의 연구도 요구되지 않는다는 것을 표시하는 일부 유사한 지정 유형)로 지정된다.

[0075] 단계(725)에서의 임계치 테스트의 결과가 포지티브라고 가정하면, 임계치-초과 잔차는 "비정상"으로 지정되고 저장된다(단계(740)에서 도시된 바와 같음). 분석의 시간 양상이 이 시점에서 도입되고, 이전의 T₂ 시간 단계들에 대한 데이터가 조사된다(단계(750)). 그 후에, 다음 단계(760)에서 결정이 이루어지는데, 여기서 T₂ 이전의 결과들 중 q%가 모두 임계치를 초과한 경우, 현재의 데이터 포인트가 센서 결함을 표시한다는 결정이 이루어지고, 센서 그 자체의 추가의 연구가 정당화된다(단계(770)). 그렇지 않고, 결정 단계(760)에서 임계치가 초과되지 않는 경우, 현재의 데이터 포인트는 "잡음있음"(또는 유사한 지정)으로 라벨링되고, 조사 프로세스가 계속된다.

[0076] 도 6에서 간략히 서술된 바와 같은 프로세스의 바람직한 실시예에서, q=100의 값이 사용되어, 특정 센서가 결함이 있을 가능성이 있는 것으로 평가되는 것을 요청하기 전에 과거 T₂ 값들 모두가 "비정상"으로서 지정되는 것이 요구된다는 것을 주목해야 한다.

[0077] **결함있는 가스 터빈 센서를 검출하는 데 활용되는 최소 코딩의 상세한 예들**

[0078] 센서 결함을 검출하기 위한 본 발명의 최소 코딩 접근법의 정확도를 결정하기 위한 실험적 접근법에서, 주어진 가스 터빈에 대한 이력 데이터의 세트가 연구되었다. 센서 데이터(MW)는, 머신이 계속 작동되고 있는지의 여부 및 판독들이 연속적인 실제 값들인지 여부를 표시하는 센서이다. 데이터는 다음과 같이 사전-프로세싱되었는데: (1) MW 출력의 최대 값 아래의 20% 미만의 값들의 명백한 이상점들 및 다른 데이터 포인트들이 제거되었고; (2) (예컨대, "ROC 미만 영역" 기준들을 사용하여) MW 판독들과 관련이 없는 리던던트 센서들(및 그들의 데이터)을 제거하기 위해 피쳐 선택(feature selection)이 수행되었고; 그리고 (3) MW와의 약한 선형 상관 관계를 나타낸 모든 나머지 센서들이 제거되었다. 궁극적으로, 15개의 센서들의 세트가 추가의 평가

를 위해 남았다.

[0079] 나머지 15개의 센서들과 연관된 데이터의 세트는 2개의 그룹들로 분할되었으며, 제 1 그룹은 트레이닝 데이터로서 사용되고, 제 2 그룹은 테스트 데이터로서 사용되었다. 트레이닝 데이터는 에러가 없는 것으로 가정되었다. 트레이닝 데이터의 랜덤하게(randomly) 선택된 위치들에서 랜덤 에러(random error)(이 경우, 증가하는 램프(increasing ramp)의 형태로)가 도입되었다.

[0080] 도 2, 도 5 및 도 6의 흐름도들에서 간략히 서술된 바와 같은 프로세스들이 데이터를 프로세싱(process)하기 위해 사용되었으며, 최소 코딩 알고리즘으로서 수학식(2)에서 사용된 $\alpha = 0.001$ 및 $\beta = 0.01$ 의 값들이 데이터에 적용되고, 잔차들이 계산되었다. "비정상" 데이터(즉, 잠재적으로 결함있는 센서)를 정의하기 위한 임계치 방법들 및 일-클래스 SVM 모델 둘 모두가 평가되었다. 일-클래스 SVM 모델의 경우, 3000개의 트레이닝 잔차들의 샘플링된 세트가 사용되어 모델을 형성하였다. SVM 모델의 시간-기반 컴포넌트의 경우, $p=50$ 의 값이 선택되었고, $T=50$ 의 값이 사용되었다(즉, 결함있는 센서의 가능성을 선언하기 위해서는 과거 50 시간 단계들에 대한 데이터의 적어도 50%가 "비정상"이도록 요구되었음). 임계치 평가 프로세스의 경우, $q=100$ 의 값이 선택되었고, $T_2=24$ 와 함께 사용되었다(즉, 현재의 데이터 포인트를 "비정상"으로서 지정하고 결함있는 센서의 가능성을 표시하기 위해서는 과거 24 시간 단계들에 대한 잔차들의 100%가 임계치를 초과해야 함).

[0081] 위에서 언급된 바와 같이, 본 발명의 현저한 양상은, 최소 코딩 프로세스들과 함께 통상적으로 사용되는 L-2 노름 대신에, L-1 노름을 사용하여 잔차들을 계산하는 것이다. 일반적으로, (행렬에서 벡터들로서 디스플레이되는(displayed)) 데이터 세트의 "노름"은 아래와 같이 정의될 수 있다:

[0082]
$$\|X\|_j = (\sum_i |x_i|^j)^{1/j}$$

[0083] L-1 노름의 경우(여기서 $j=1$), 이 관계식은 아래와 같이 단순화되고:

[0084]
$$\|X\|_1 = (\sum_i |x_i|)$$

[0085] 그리고 종종 "최소 절대 편차(least absolute deviation)"(LAD) 또는 "최소 절대 에러(least absolute error)"(LAE)로 지칭된다. 이와 비교하여, L-2는 아래와 같이 종래의 유클리드 "거리" 측정(Euclidean "distance" measure)으로서 표현될 수 있다:

[0086]
$$\|X\|_2 = (\sum_i |x_i|^2)^{1/2} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

[0087] 가스 터빈 센서들에 의해 수집된 데이터가 상대적으로 큰 잡음 컴포넌트를 포함할 수 있기 때문에, L-2 노름보다 L-1 노름이 이러한 원하지 않는 정보에 덜 민감할 것이라고 결정되었다. 비교를 위해, L-2 노름을 사용하여 계산들의 세트가 또한 준비되었다. 위에서 언급된 도 3 및 도 4에 대한 참조가 이제 이루어지는데, 여기서 이들 2개의 플롯(plot)들은 L-1 노름을 사용할 때 획득된 결과들의 개선을 예시한다. 도 3은 L-2 잔차들로서 상대적으로 많은 수의 스파이크(spike)들을 도시하는, 최소 코딩 프로세스의 결과들의 플롯이다. 이에 반해, (본 발명에 따라 L-1 노름을 사용하는 것에 기반한) 도 4의 플롯은 훨씬 더 적은 수의 잔차들을 포함한다. 명백하게, L-2 노름은 더 많은 스파이크들을 생성하는데, 이는 실제 센서 판독들에서 작은 편차들을 검출하는 능력이 검출을 하는 것을 더 어렵게 만들며, 여기서 이러한 미묘한 데이터는 센서 고장을 더 많이 표시할 수 있고 본 발명의 목적들에 있어서 더 중요할 수 있다.

[0088] 또한 비교를 위해, "비정상" 데이터의 최종 결정을 하는 시간 컴포넌트를 무시한 결과들의 세트가 생성되었다. 아래의 테이블에서, 이 결과들의 세트는 "베이스라인 알고리즘(baseline algorithm)"으로 식별된다. 테이블에서, 잔차들(비정상성들)은 "+"로 표기되고, 정상 데이터는 "-"로 표기된다. 특정 태블러 형태(tabular form)는, 정상 데이터가 비정상성으로서 잘못 분류된 경우뿐만 아니라 비정상성이 정상 데이터로 잘못 분류된 경우의 인스턴스(instance)들의 형태로 "허위 포지티브들"을 (헤더(header)들 "예측" 및 "실제"를 사용하여) 보여준다.

L-1 노름 : 랜덤 포인트들	베이스라인 알고리즘		SVM 모델		임계치	
	실제 예측	+	-	+	-	+
+	571	1094	582	232	456	18
-	293	140254	282	141116	408	141330
(포인트 51010)						
+	218	1271	205	413	167	150
-	646	140077	659	140935	697	141198
(포인트 23189)						
+	279	1268	270	413	254	150
-	585	140080	594	140935	610	141198
(포인트 89904)						
+	355	1278	345	413	247	150
-	509	140070	519	140935	617	141198
(포인트 56385)						
+	369	1269	363	413	220	150
-	495	140079	501	140935	644	141198
(포인트 78503)						

[0089]

[0090]

위의 테이블의 조사로부터, 본 발명의 접근법들(SVM 모델 및 임계치) 둘 모두가, 베이스라인 접근법과 비교할 때 (더 낮은 수들의 허위-포지티브들의 관점들에서) 개선된 결과들을 제공한다는 것이 명백하다. (L-2 및 L-1 노름 둘 모두를 사용한 각각의 인스턴스에서 계산된 값들을 포함하여) 총 10개의 랜덤 시간 포인트들에 대한 더 완전한 세트의 결과들이 본 명세서의 부록(appendix)에서 보여진다. 부록에 포함된 완전한 세트의 결과들은 추가로, 각각의 인스턴스에서 L-2 노름 대신에 L-1 노름의 사용이 더 정확한 결과를 제공한다는 것을 보여주며, 그에 따라 L-1 노름의 선택이 더 양호한 연산임을 입증한다.

[0091]

위에서 설명된 방법은 위에서 설명된 바와 같이 컴퓨터(computer)에 의해 실행되는 프로그램 모듈(program module)들에 의해 구현될 수 있다. 일반적으로, 프로그램 모듈들은, 특정 태스크(task)들을 수행하거나 또는 특정한 추상적 데이터 유형들을 구현하는 루틴(routine)들, 오브젝트(object)들, 컴포넌트들, 데이터 구조들 등을 포함한다. 본원에서 사용되는 바와 같은 "프로그램"이라는 용어는 단일 프로그램 모듈 또는 협력하여 동작하는 다수의 프로그램 모듈들을 함축할 수 있다. 본 개시내용은, 개인 컴퓨터(PC; personal computer)들, 핸드-헬드 디바이스(hand-held device)들, 멀티-프로세서 시스템(multi-processor system)들, 마이크로프로세서-기반 프로그램가능 가전제품(microprocessor-based programmable consumer electronics), 네트워크 PC(network PC)들, 미니-컴퓨터(mini-computer)들, 메인프레임 컴퓨터(mainframe computer)들 등을 포함한 다양한 유형들의 컴퓨터들 상에서 구현될 수 있다. 본 개시내용은 또한, 통신 네트워크를 통해 링크된(linked) 원격 프로세싱 디바이스(remote processing devices)들에 의해 태스크들이 수행되는 분산형 컴퓨팅 환경들에서 이용될 수 있다. 분산형 컴퓨팅 환경에서, 모듈들은 로컬(local) 및 원격 메모리 저장 디바이스들 둘 모두에 위치될 수 있다.

[0092]

위의 방법론을 구현하기 위한 예시적인 프로세싱 모듈은 하드와이어되거나(hardwired) 또는 개별 메모리에 저장될 수 있으며, 개별 메모리는, ROM 또는 다른 유형의 하드 자기 드라이브(hard magnetic drive), 광학 저장소, 테이프(tape) 또는 플래시 메모리(flash memory)와 같은 컴퓨터 판독가능 매체로부터 프로세서 또는 복수의 프로세서들의 메인 메모리(main memory)로 판독된다. 메모리 매체들에 저장된 프로그램의 경우, 모듈에서의 명령들의 시퀀스(sequence)들의 실행은 프로세서로 하여금 본원에서 설명된 프로세스 단계들을 수행하게 한다. 본 개시내용의 실시예들은 하드웨어와 소프트웨어의 임의의 특정 조합으로 제한되지 않으며, 기술한 내용을 구현하기 위해 요구되는 컴퓨터 프로그램 코드는 당업자에 의해 개발될 수 있다.

[0093]

본원에서 이용되는 바와 같은 "컴퓨터-판독가능 매체"라는 용어는, 명령들을 하나 또는 그 초과수의 프로세서들에

제공하는 또는 제공하는 데 참여하는 임의의 유형적 머신-인코딩된 매체(tangible machine-encoded medium)를 의미한다. 예컨대, 컴퓨터-판독가능 매체는 하나 또는 그 초과와 광학 또는 자기 메모리 디스크(disk)들, 플래시 드라이브(flash drive)들 및 카드(card)들, 판독-전용 메모리, 또는 통상적으로 메인 메모리를 구성하는 DRAM과 같은 랜덤 액세스 메모리(random access memory)일 수 있다. 이러한 매체들은 유형적이지 않은 전파된 신호들을 제외한다. 캐시된(cached) 정보는 컴퓨터-판독가능 매체 상에 저장되는 것으로 고려된다. 컴퓨터-판독가능 매체들의 통상의 수단들은 당해 기술 분야에 잘 알려져 있으며, 여기서 상세하게 설명될 필요가 없다.

[0094] 전술한 상세한 설명은 모든 관점에서 제한적인 것이 아니라 예시적이고 대표적인 것으로서 이해되어야 하며, 본원의 본 개시내용의 범위는 상세한 설명으로부터 결정되는 것이 아니라, 특허법에 의해 허용되는 최대 범위에 따라 해석되는 청구항들로부터 결정될 것이다. 본 개시내용의 범위 및 사상으로부터 벗어남이 없이 다양한 수정들이 당업자들에 의해 구현될 수 있다는 것이 이해되어야 한다.

[0095] **부록 - 잔차에 대한 영향; L-2 노름을 L-1 노름과 비교**

시간 인덱스 51010 에서 선택된 랜덤 포인트

L-2 노름		베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-	
+	443	1130	431	499	462	64	
-	421	140218	433	140849	402	141284	
정밀도	0.281627		0.463441		0.878327		
리콜 (Recall)	0.512731		0.498843		0.534722		
L-1 노름		베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-	
+	571	1094	582	232	456	18	
-	293	140254	282	141116	408	141330	
정밀도	0.342943		0.714988		0.962025		
리콜	0.66088		0.673611		0.527778		

[0096]

시간 인덱스 23189 에서 선택된 랜덤 포인트

L-2 노름		베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-	
+	77	1345	68	717	167	196	
-	787	140003	796	140631	697	141152	
정밀도	0.054149		0.086624		0.460055		
리콜	0.08912		0.078704		0.193287		
L-1 노름		베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-	
+	218	1271	205	413	167	150	
-	646	140077	659	140935	697	141198	
정밀도	0.146407		0.331715		0.526814		
리콜	0.252315		0.237269		0.193287		

[0097]

시간 인덱스 76725 에서 선택된 랜덤 포인트

L-2 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	193	1337	177	719	224	196
-	671	140011	687	140629	640	141152
정밀도	0.126144		0.197545		0.533333	
리콜	0.22338		0.204861		0.259259	

L-1 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	334	1264	317	413	224	150
-	530	140084	547	140935	640	141198
정밀도	0.209011		0.434247		0.59893	
리콜	0.386574		0.366898		0.259259	

[0098]

시간 인덱스 80117 에서 선택된 랜덤 포인트

L-2 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	181	1333	167	719	169	196
-	683	140015	697	140629	695	141152
정밀도	0.119551		0.188488		0.463014	
리콜	0.209491		0.193287		0.195602	

L-1 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	369	1264	357	413	168	150
-	495	140084	507	140935	696	141198
정밀도	0.225964		0.463636		0.528302	
리콜	0.427083		0.413194		0.194444	

[0099]

시간 인덱스 89904 에서 선택된 랜덤 포인트

L-2 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	114	1336	94	719	257	196
-	750	140012	770	140629	607	141152
정밀도	0.078621		0.115621		0.567329	
리콜	0.131944		0.108796		0.297454	

L-1 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	279	1268	270	413	254	150
-	585	140080	594	140935	610	141198
정밀도	0.180349		0.395315		0.628713	
리콜	0.322917		0.3125		0.293981	

[0100]

시간 인덱스 56385 에서 선택된 랜덤 포인트

L-2 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	178	1348	158	719	247	196
-	686	140000	706	140629	617	141152
정밀도	0.116645		0.18016		0.557562	
리콜	0.206019		0.18287		0.28588	
L-1 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	355	1278	345	413	247	150
-	509	140070	519	140935	617	141198
정밀도	0.217391		0.455145		0.622166	
리콜	0.41088		0.399306		0.28588	

[0101]

시간 인덱스 78503 에서 선택된 랜덤 포인트

L-2 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	227	1336	231	719	220	196
-	637	140012	633	140629	644	141152
정밀도	0.145234		0.243158		0.528846	
리콜	0.262731		0.267631		0.25463	
L-1 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	369	1269	363	413	220	150
-	495	140079	501	140935	644	141198
정밀도	0.225275		0.467784		0.594595	
리콜	0.427083		0.420139		0.25463	

[0102]

시간 인덱스 56711 에서 선택된 랜덤 포인트

L-2 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	184	1348	143	719	412	196
-	680	140000	721	140629	452	141152
정밀도	0.120104		0.165893		0.677632	
리콜	0.212963		0.165509		0.476852	
L-1 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	406	1277	408	413	412	150
-	458	140071	456	140935	452	141198
정밀도	0.241236		0.496955		0.733096	
리콜	0.469907		0.472222		0.476852	

[0103]

시간 인덱스 31449 에서 선택된 랜덤 포인트

L-2 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	129	1348	117	719	238	196
-	735	140000	747	140629	626	141152
정밀도	0.087339		0.139952		0.548387	
리콜	0.149306		0.135417		0.275463	
L-1 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	253	1278	237	413	238	150
-	611	140070	627	140935	626	141198
정밀도	0.165251		0.364615		0.613402	
리콜	0.292824		0.274306		0.275463	

[0104]

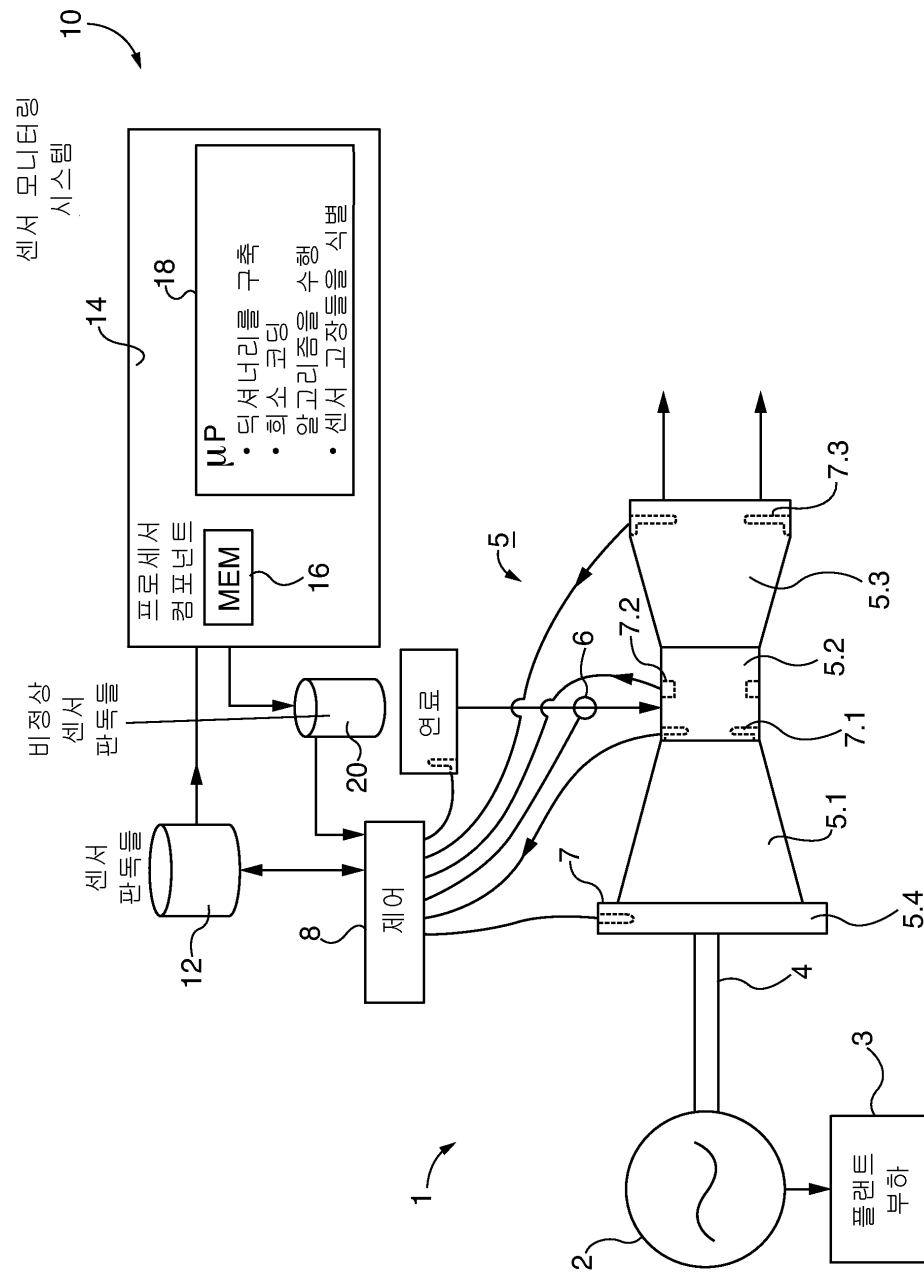
시간 인덱스 102457 에서 선택된 랜덤 포인트

L-2 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	194	1343	180	719	229	196
-	670	140005	684	140629	635	141152
정밀도	0.12622		0.200222		0.538824	
리콜	0.224537		0.208333		0.265046	
L-1 노름	베이스라인 알고리즘		SVM		임계치	
예측 \ 실제	+	-	+	-	+	-
+	349	1270	343	413	229	150
-	515	140078	521	140935	635	141198
정밀도	0.215565		0.453704		0.604222	
리콜	0.403935		0.396991		0.265046	

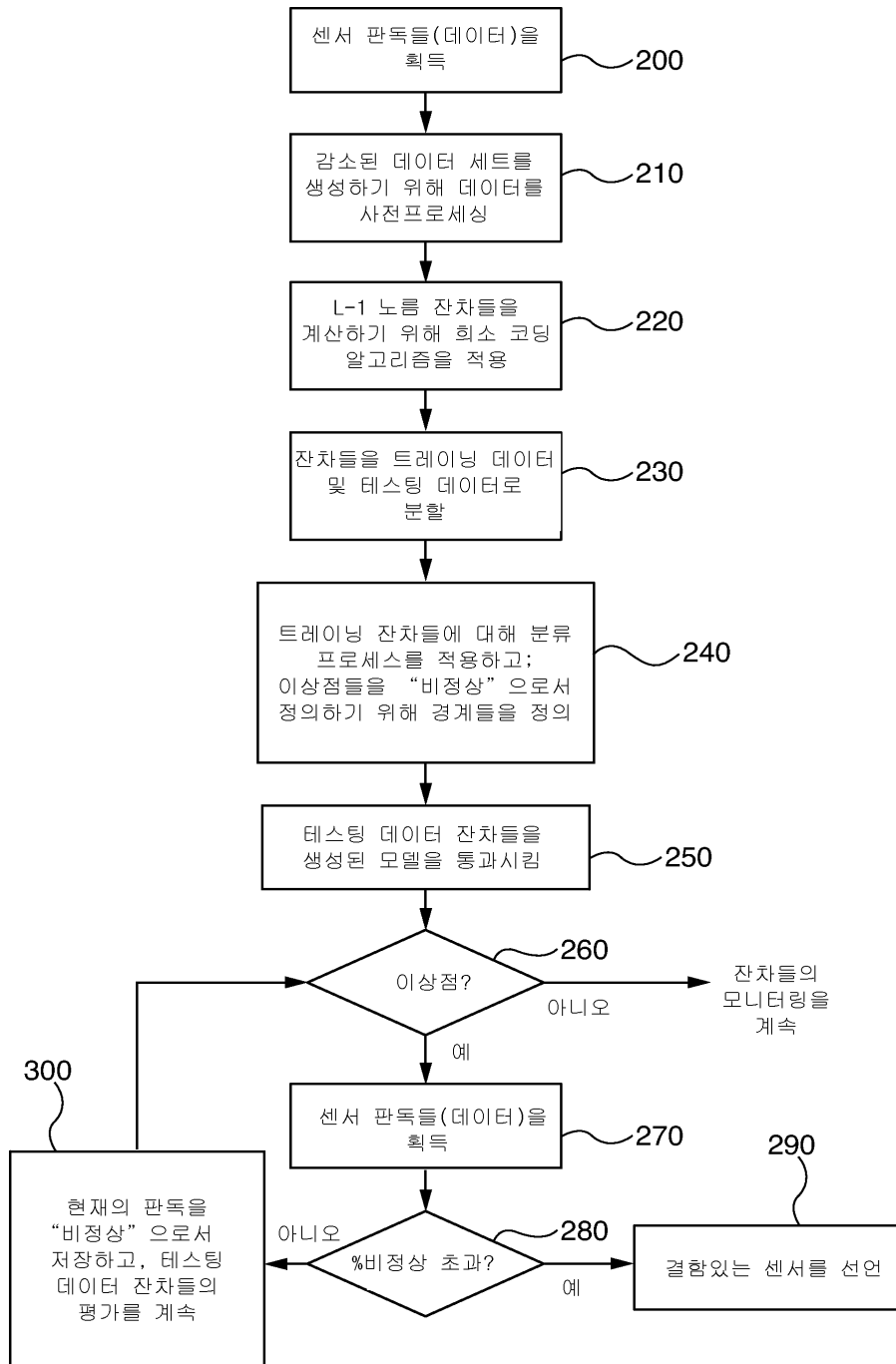
[0105]

도면

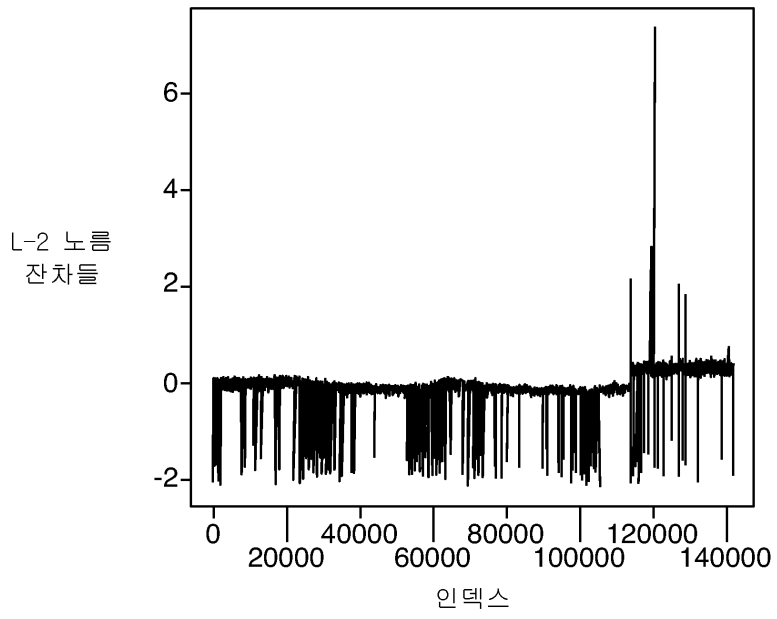
도면1



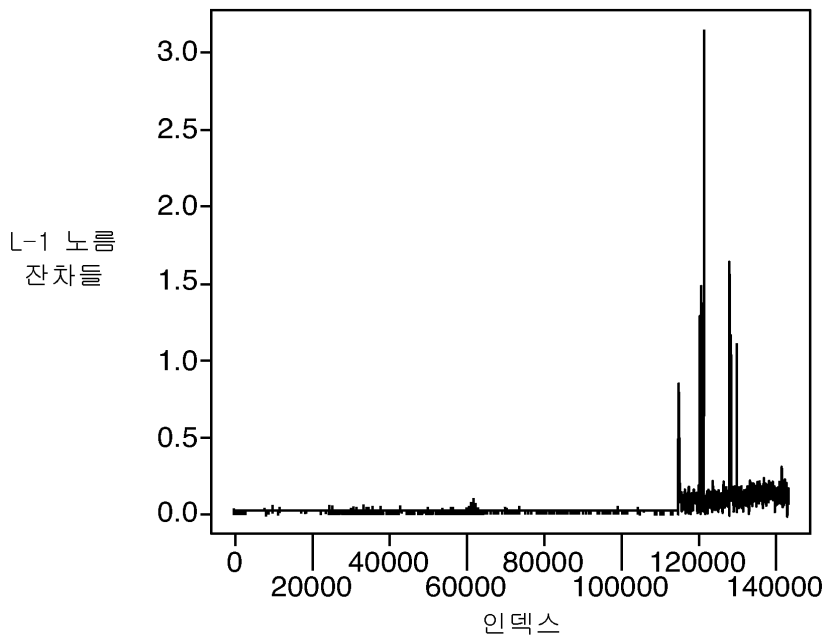
도면2



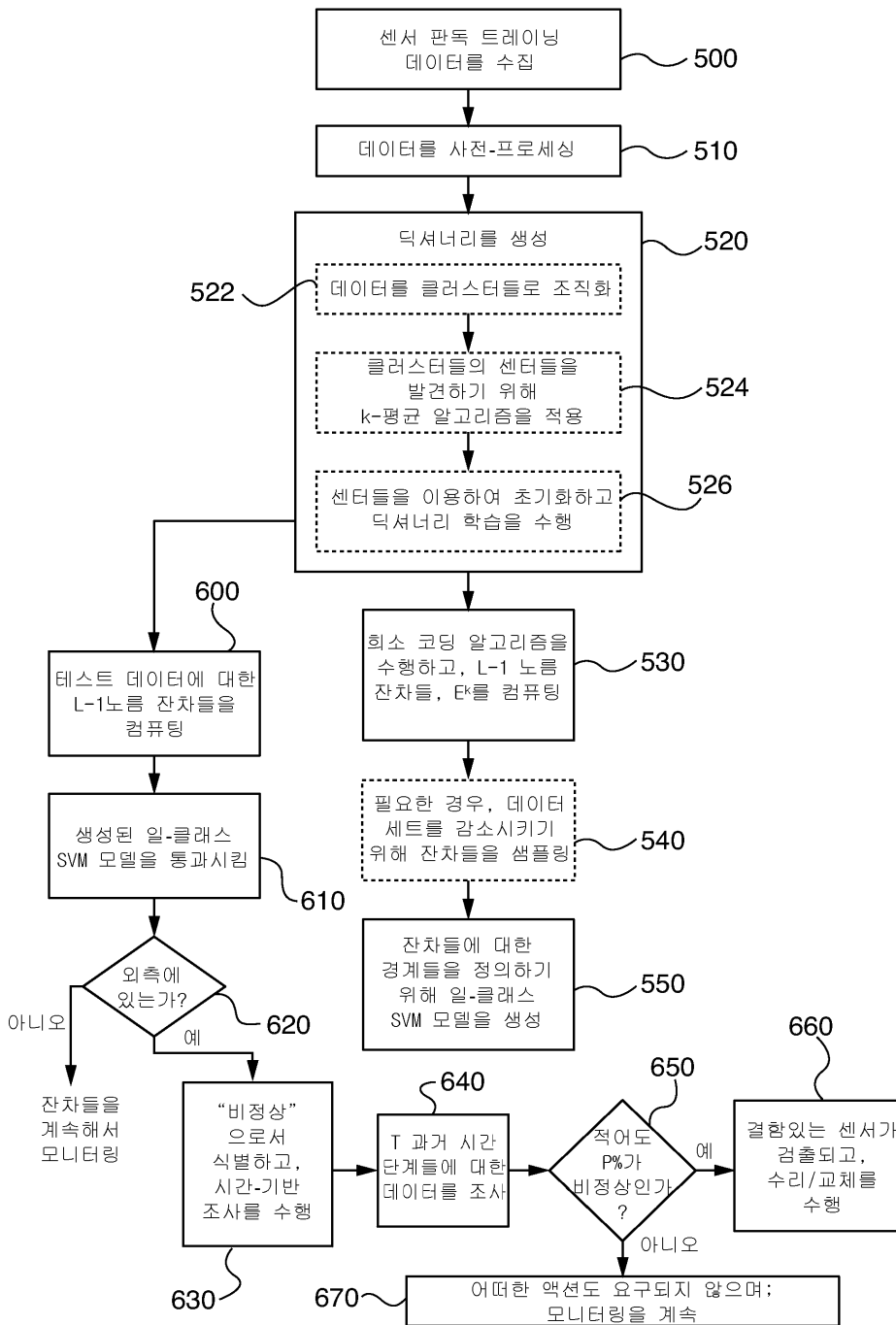
도면3



도면4



도면5



도면6

