



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년12월04일
(11) 등록번호 10-2186852
(24) 등록일자 2020년11월30일

(51) 국제특허분류(Int. Cl.)
G16B 45/00 (2019.01) G16B 40/00 (2019.01)
(52) CPC특허분류
G16B 45/00 (2019.02)
G16B 40/00 (2019.02)
(21) 출원번호 10-2018-0071480
(22) 출원일자 2018년06월21일
심사청구일자 2018년06월21일
(65) 공개번호 10-2019-0143677
(43) 공개일자 2019년12월31일
(56) 선행기술조사문헌
JP2007086923 A
(뒷면에 계속)

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
한병찬
서울특별시 서대문구 연희로32길 48, 108동 203호 (연희동, 연희동성원아파트)
노승효
대전광역시 서구 복수서로 50, 101동 311호 (복수동, 삼익목화아파트)
강준희
서울특별시 양천구 목동중앙서로7다길 46, 302호 (목동)
(74) 대리인
특허법인 하나

전체 청구항 수 : 총 7 항

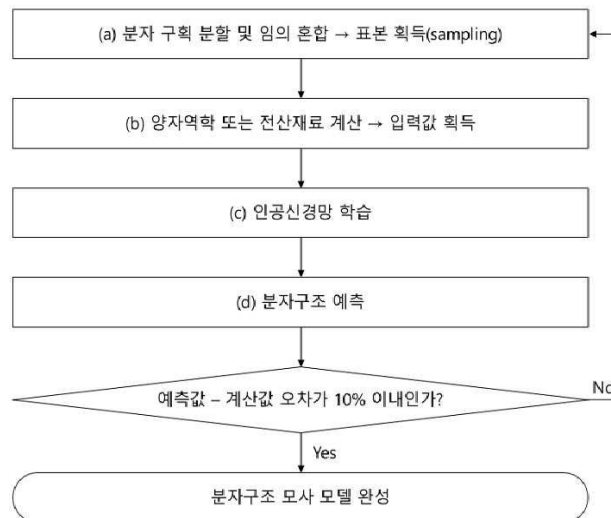
심사관 : 이재영

(54) 발명의 명칭 분자구조의 모사방법

(57) 요약

본 발명의 일 측면은, (a) 분자를 복수의 원자를 포함하는 복수의 구획으로 분할하고, 상기 복수의 구획 중 2 이상을 임의로 혼합하여 표본을 얻는 단계; (b) 상기 표본에 대한 양자역학 또는 전산재료 계산으로 입력값을 얻는 단계; (c) 상기 입력값 중 일부를 이용하여 인공신경망을 학습시키는 단계; 및 (d) 상기 인공신경망을 이용하여 상기 분자 중 상기 표본을 제외한 부분의 구조를 예측하는 단계;를 포함하고, 상기 (a) 단계에서 상호 인접한 각각의 구획 중 일부가 중첩되도록 분할하는 것을 특징으로 하는 분자구조의 모사방법을 제공한다.

대표도 - 도2



(56) 선행기술조사문헌
 KR1020080099559 A
 KR1020120085178 A
 KR1020140145753 A
 US20140079746 A1
 W02018220368 A1

이 발명을 지원한 국가연구개발사업

과제고유번호	2013M3A6B1078882
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	글로벌프론티어사업
연구과제명	[Ezbaro] (총괄/0세부)Materials Computation (3단계)(1/5)
기 여 율	1/2
과제수행기관명	연세대학교
연구기간	2018.03.01 ~ 2018.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	20173010032080
부처명	산업통상자원부
과제관리(전문)기관명	한국에너지기술평가원
연구사업명	에너지기술개발사업
연구과제명	[RCMS] 한국에너지기술연구원/3상분리형 촉매층 기반 연료전지 성능 및 내구성 향상
기술 개발(2/4)	
기 여 율	1/2
과제수행기관명	연세대학교
연구기간	2018.01.01 ~ 2018.12.31

명세서

청구범위

청구항 1

- (a) 분자를 복수의 원자를 포함하는 복수의 구획으로 분할하고, 상기 복수의 구획 중 2 이상을 임의로 혼합하여 표본을 얻는 단계;
- (b) 상기 표본에 대한 양자역학 계산으로 입력값을 얻는 단계;
- (c) 상기 입력값 중 일부를 이용하여 인공신경망을 학습시키는 단계; 및
- (d) 상기 인공신경망을 이용하여 상기 분자 중 상기 표본을 제외한 부분의 구조를 예측하는 단계;를 포함하고, 상기 (a) 단계에서 상호 인접한 각각의 구획 중 일부가 중첩되도록 분할하는 것을 특징으로 하는 분자구조의 모사방법.

청구항 2

제1항에 있어서,
상기 구획은 20개 이하의 원자를 포함하는 것을 특징으로 하는 분자구조의 모사방법.

청구항 3

제1항에 있어서,
상기 상호 인접한 각각의 구획에 포함된 총 원자 수의 20~30%가 중첩되는 것을 특징으로 하는 분자구조의 모사방법.

청구항 4

제1항에 있어서,
상기 분자 중 상기 표본의 비율은 0.01% 이하인 것을 특징으로 하는 분자구조의 모사방법.

청구항 5

제1항에 있어서,
상기 입력값 중 10~50%를 이용하여 인공신경망을 학습시키는 것을 특징으로 하는 분자구조의 모사방법.

청구항 6

제1항에 있어서,
상기 입력값은 상기 구획 내에 포함된 원자의 결합, 각도, 면각도 및 이들 중 2 이상의 조합으로 이루어진 군에서 선택된 하나인 것을 특징으로 하는 분자구조의 모사방법.

청구항 7

제1항에 있어서,
상기 (d) 단계에서 상기 분자 중 상기 표본을 제외한 부분에 대해 예측된 구조와 양자역학 계산값 간의 오차가 10% 이내일 때까지 상기 (a) 내지 (d) 단계를 반복하는 것을 특징으로 하는 분자구조의 모사방법.

발명의 설명

기술 분야

[0001] 본 발명은 분자구조의 모사방법에 관한 것으로, 더 상세하게는 인공신경망을 이용하여 분자 내(intra-

molecular) 및 분자 간(inter-molecular) 상호작용을 정확하게 모사함으로써 분자구조를 고속으로 예측하는 방법에 관한 것이다.

배경 기술

- [0002] 고분자, 거대분자, 화합물과 같은 다양한 물성의 정확한 이해는 그 분자의 용도의 타당성을 검토하거나, 합성 및 정제과정을 설계하고 보관, 운반, 사용, 폐기의 방법과 조건을 설정하는 등 생산과 소비의 전 과정에 걸친 제반 의사결정 사항들에 결정적이므로 산업적, 학문적으로 매우 중요한 문제이다.
- [0003] 분자의 물성을 가장 직접적으로 획득할 수 방법은, 실험을 이용하는 것이나, 정제된 시료의 준비, 정확한 측정을 위한 환경의 구축 등 여러 가지 측면에서 상당한 비용과 시간이 요구되며 경우에 따라서는 불가능할 수도 있다. 따라서, 모델링이나 전산모사 등 다양한 대안들이 제안되어왔다.
- [0004] 예측모형을 구축하기 위한 방법으로 양자역학에 기반한 제일원리 전산모사는 분자의 물성 예측에 매우 높은 정확도를 가지는 것으로 알려져 있다. 하지만, 일반적으로 제일원리 전산모사법은 분자의 크기 증가와 구조의 복잡성에 따라 컴퓨팅의 로드(load)가 매우 높다는 단점이 있다. 따라서, 적절한 분자구조의 샘플링과 계산 난이도를 적용한 후, 예상되는 오차를 QSPR(quantitative structure-property relationship)법을 이용하여 정밀하게 보정하는 방법을 선택하였다. QSPR은 기본적으로 주어진 분자의 물성은 그 구조적 특성의 함수라는 가정에서 출발하며, 다양한 분자제어 인자(molecular control descriptor)를 발견하는 것이 그 핵심이다. 현재까지 제안된 분자제어 인자들의 종류는 수 천에 이르며, 한 분자 내의 원자의 수와 같은 단순한 것들로부터 분자의 모양이나 연결상태, 전기화학적 특성과 같은 복잡한 것들에 이르기까지 매우 다양하다.
- [0005] QSPR 모형을 만드는 또 다른 방법은 인공신경망(neural network)을 이용하는 것이다. 인공신경망 기법은 인간의 신경세포의 작동기제를 모방하는 기계적 학습법을 창조하려는 것으로, 20세기 중반에 등장한 이래, 다양한 분야의 정보처리에 활용되는 기술이다.
- [0006] 도 1은 인공신경망의 전형적인 한 예를 보여준다. 인공신경망에는 입력 데이터를 수용하는 입력층(input layer)과 출력데이터를 만드는 출력층(output layer), 이들 사이에 위치한 은닉층(hidden layer)이 존재하며 각 층은 하나 이상의 노드(node)들로 구성되어 있다. 은닉층의 각 노드들은 입력층과 출력층의 노드들과 가중치(weight)라 불리는 양(w_{ij} , w'_{ij})으로 상호 연결되어 있다. 은닉층과 출력층의 각 노드들은 이전 단계의 노드들로부터 입력값을 받은 뒤 이를 연산하여 출력값을 만드는데, 이 때, 활성화 함수(activation function)라 불리는 함수(f_1 , f_2)를 적용한다. 이러한 인공신경망 이론을 실제 정보처리 기술에 적용하기 위해서는 다양한 입력값과 이에 대응하는 출력값을 함께 묶어 놓은 샘플, 즉, 훈련 집합을 이용하여 컴퓨터를 학습시키는 과정이 필요하다. 구체적으로, 주어진 입력에 대한 인공신경망의 예측값과 실제값의 차이가 최소가 되도록 역전파(back propagation) 알고리즘을 사용하여 각 연결점에 대한 가중치를 최적화하는 것이다. 일단 이러한 학습을 마친 인공신경망은 문제해결에 필요한 규칙이나 지식을 별도로 제공하지 않아도 자율적으로 규칙을 수립을 통해 임의의 입력값에 대해서도 타당성있는 출력값을 도출함으로써, 미지 화합물의 물성 예측과 같은 기반 이론이 결여된 시스템에 매우 유용한 수단으로 이용되고 있다.
- [0007] 다만, 양자역학 기반의 순수 제일원리 전산모사는 정확도가 높은 반면에, 약 1,000개 이상의 원자로 구성된 분자시스템에는 그 적용이 사실상 불가능하고, 특히, 고분자 내에서의 상호작용, 예를 들어, 접힘(folding)과 고분자 간 상호작용 모사에 매우 취약하다. 또한, 인공신경망 방법은 계산의 정확도와 속도를 모두 개선할 수 있으나, 이 또한 분자 내 및 분자 간 상호작용을 위한 훈련 집합을 구축하기 어려운 문제가 있다.

발명의 내용

해결하려는 과제

- [0008] 본 발명은 전술한 종래 기술의 문제점을 해결하기 위한 것으로, 본 발명의 목적은 인공신경망을 이용하여 상기 분자의 거동, 즉, 분자 내(intra-molecular) 및 분자 간(inter-molecular) 상호작용을 신속하고 정확하게 예측, 모사할 수 있는 분자구조의 모사방법을 제공하는 것이다.

과제의 해결 수단

- [0009] 본 발명의 일 측면은, (a) 분자를 복수의 원자를 포함하는 복수의 구획으로 분할하고, 상기 복수의 구획 중 2

이상을 임의로 혼합하여 표본을 얻는 단계; (b) 상기 표본에 대한 양자역학 또는 전산재료 계산으로 입력값을 얻는 단계; (c) 상기 입력값 중 일부를 이용하여 인공신경망을 학습시키는 단계; 및 (d) 상기 인공신경망을 이용하여 상기 분자 중 상기 표본을 제외한 부분의 구조를 예측하는 단계;를 포함하고, 상기 (a) 단계에서 상호 인접한 각각의 구획 중 일부가 중첩되도록 분할하는 것을 특징으로 하는 분자구조의 모사방법을 제공한다.

- [0010] 일 실시예에 있어서, 상기 구획은 20개 이하의 원자를 포함할 수 있다.
- [0011] 일 실시예에 있어서, 상기 상호 인접한 각각의 구획에 포함된 총 원자 수의 20~30%가 중첩될 수 있다.
- [0012] 일 실시예에 있어서, 상기 분자 중 상기 표본의 비율은 0.01% 이하일 수 있다.
- [0013] 일 실시예에 있어서, 상기 입력값 중 10~50%를 이용하여 인공신경망을 학습시킬 수 있다.
- [0014] 일 실시예에 있어서, 상기 입력값은 상기 구획 내에 포함된 원자의 결합, 각도, 면각도 및 이들 중 2 이상의 조합으로 이루어진 군에서 선택된 하나일 수 있다.
- [0015] 일 실시예에 있어서, 상기 (d) 단계에서 상기 분자 중 상기 표본을 제외한 부분에 대해 예측된 구조와 양자역학 또는 전산재료 계산값 간의 오차가 10% 이내일 때까지 상기 (a) 내지 (d) 단계를 반복할 수 있다.

발명의 효과

- [0016] 본 발명의 일 측면에 따른 분자구조의 모사방법은, 분자를 복수의 구획으로 구조적으로 분할하고 이들 구획을 임의로 혼합하여 획득된 표본으로부터 인공신경망에 대한 입력값을 도출하고, 상기 인공신경망을 이용하여 상기 분자의 거동, 즉, 분자 내(intra-molecular) 및 분자 간(inter-molecular) 상호작용을 신속하고 정확하게 예측, 모사할 수 있다.
- [0017] 본 발명의 효과는 상기한 효과로 한정되는 것은 아니며, 본 발명의 상세한 설명 또는 청구범위에 기재된 발명의 구성으로부터 추론 가능한 모든 효과를 포함하는 것으로 이해되어야 한다.

도면의 간단한 설명

- [0018] 도 1은 본 발명의 일 실시예에 따른 인공신경망을 도식화한 것이고;
- 도 2는 본 발명의 일 실시예에 따른 분자구조의 모사방법을 도식화한 것이고;
- 도 3은 본 발명의 일 실시예에 따른 분자의 상호작용을 도식화한 것이고;
- 도 4는 본 발명의 일 실시예에 따른 분자의 구획 분할 및 임의의 혼합을 도식화한 것이고;
- 도 5는 본 발명의 일 실시예에 따른 분자구조의 모사결과를 도식화한 것이고;
- 도 6은 본 발명의 일 실시예에 따른 분자구조의 모사방법의 정확도를 평가한 결과이다.

발명을 실시하기 위한 구체적인 내용

- [0019] 이하에서는 첨부한 도면을 참조하여 본 발명을 설명하기로 한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 따라서 여기에서 설명하는 실시예로 한정되는 것은 아니다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0020] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 부재를 사이에 두고 "간접적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 구비할 수 있다는 것을 의미한다.
- [0021] 이하, 첨부된 도면을 참고하여 본 발명의 실시예를 상세히 설명하기로 한다.
- [0022] 분자구조의 모사방법
- [0023] 도 2는 본 발명의 일 실시예에 따른 분자구조의 모사방법을 도식화한 것이다. 도 2를 참고하면, 본 발명의 일 실시예에 따른 분자구조의 모사방법은, (a) 분자를 복수의 원자를 포함하는 복수의 구획으로 분할하고, 상기 복수의 구획 중 2 이상을 임의로 혼합하여 표본을 얻는 단계; (b) 상기 표본에 대한 양자역학 계산으로 입력값을

얻는 단계; (c) 상기 입력값 중 일부를 이용하여 인공신경망을 학습시키는 단계; 및 (d) 상기 인공신경망을 이용하여 상기 분자 중 상기 표본을 제외한 부분의 구조를 예측하는 단계;를 포함할 수 있다.

- [0024] 상기 (a) 단계에서는 분자를 복수의 원자를 포함하는 복수의 구획으로 분할하고, 상기 복수의 구획 중 2 이상을 임의로 혼합하여 표본을 얻을 수 있다. 상기 분자는 고분자, 거대분자, 또는 화합물로서 1,000개 이상의 원자를 포함하는 것일 수 있고, 이 때, 상기 분자의 구조를 정확하게 예측, 모사하기 위해서는 분자 내 상호작용뿐만 아니라 분자 간 상호작용도 고려할 필요가 있다.
- [0025] 도 3은 본 발명의 일 실시예에 따른 분자의 상호작용을 도식화한 것으로서, 도 3(a) 및 3(b)는 각각 분자 내 및 분자 간 상호작용을 나타낸다. 도 3을 참고하면, 상기 분자 내 상호작용, 예를 들어, 분자의 접힘(folding)은 단일 분자 내에 존재하는 전자 간의 상호작용에 기인한 것이므로, 상기 구획이 불연속적으로, 즉, 상호 중첩없이 분할되면 이들 간의 상기 분자를 구성하는 인접 구획 간의 상호작용을 정확하게 학습, 모사, 예측하기 어렵다. 따라서, 상기 (a) 단계에서 상호 인접한 각각의 구획 중 일부가 중첩되도록 분할할 수 있다. 상기 상호 인접한 각각의 구획에 포함된 총 원자 수의 20~30%가 중첩될 수 있다. 상기 중첩된 원자의 비율이 20% 미만이면 분자 내 상호작용을 정확하게 학습, 모사, 예측하기 어렵고, 30% 초과이면 후속되는 인공신경망을 학습시키는 단계에서 과적합(overfitting)이 발생할 수 있다.
- [0026] 단일의 구획은 원자의 결합, 각도, 면각도 및 이들 중 2 이상의 조합으로 이루어진 군에서 선택된 하나, 바람직하게는, 결합, 각도 및 면각도의 조합이 모두 고려될 수 있는 최소의 단위로서, 20개 이하, 바람직하게는, 5~20개의 원자를 포함할 수 있다. 단일의 구획에 포함된 원자 수가 20개 초과이면 후속되는 양자역학 계산에 대한 부담이 가중될 수 있고, 인공신경망을 학습시키는 단계에서 과적합(overfitting)이 발생할 수 있다.
- [0027] 또한, 상기 분자 중 상기 표본의 비율은 0.01% 이하일 수 있다. 상기 표본의 비율이 0.01% 초과이면 후속되는 양자역학 계산에 대한 부담이 가중될 수 있고, 인공신경망을 학습시키는 단계에서 과적합(overfitting)이 발생하여 계산, 예측, 모사의 신속성, 정확도가 저하될 수 있다.
- [0028] 도 4는 본 발명의 일 실시예에 따른 분자의 구획 분할 및 임의 혼합을 도식화한 것이고, 도 5는 이러한 구획 및 표본을 이용하여 양자역학 계산 및 인공신경망 방법으로 모사한 분자구조를 도식화한 것이다. 도 4(a)를 참고하면, 공지의 불소계 고분자인 NafionTM 1100은 탄소(C), 산소(O), 불소(F) 및 황(S)를 포함하는 구획 1, 탄소(C), 산소(O) 및 불소(F)를 포함하는 구획 2, 및 탄소(C) 및 불소(F)를 포함하는 구획 3으로 분할될 수 있고, 구획1과 구획 2, 및 구획 2와 구획 3 사이에는 일정 수준의 중첩이 허용될 수 있다. 또한, 도 4(b)를 참고하면, 상기 구획 1 내지 구획 3 중 하나 이상을 임의로 혼합하여 하나 이상, 바람직하게는, 복수의 표본을 얻을 수 있다.
- [0029] 상기 (b) 단계에서는 상기 표본에 대한 양자역학 계산으로 인공신경망에 대한 입력값을 얻을 수 있다. 분자의 전자구조를 계산하기 위해서는 보통 순이론인 방법으로 슈뢰딩거(Schrodinger) 방정식을 풀어 전자에너지에 대한 해를 구하게 되지만 전자가 많은 계의 경우 전자상관관계(electron correlation)를 무시한 근사법을 적용한 하트리-포크(Hartree-Fock, HF) 방법[C. C. J. Roothan, Rev. Mod. Phys. 23, 69 (1951)]을 사용하여 해를 풀게 된다. 이런 근사법으로 인해 계산된 결과에서 근본적인 오차가 유발되어 다차원의 이론적인 섭동항을 추가한 포스트 하트리-포크(Post Hartree-Fock) 방법[C. Moller and M. S. Plesset, Phys. Rev. 46, 618 (1934)]들을 사용하여 더 정확한 해를 구하긴 하지만 상대적으로 많은 계산량이 요구된다. 이런 방식으로는 큰 분자를 계산하기에는 비용이나 시간의 측면에서 무리가 있다.
- [0030] 또한, 하트리-포크와 포스트 하트리-포크를 조합한 가우시안 방법[L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. 94, 7221 (1991); L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, J. Chem. Phys. 109, 7764 (1998)]은 에너지 예측에 있어 아주 적은 오차를 보이지만 여러 포스트 하트리-포크 방법에 대한 에너지 계산을 수행하기 때문에 더 많은 계산량이 요구된다.
- [0031] 많은 전자로 이루어진 분자에 대한 전자들간의 상호작용, 상관관계를 고려하기 위해 다차원의 섭동항이 추가된 파동함수 대신 전자밀도함수를 써서 총에너지의 범함수를 이용해서 바닥상태를 구하는 밀도범함수 이론(density functional theory)[R. Seeger and J. A. Pople, J. Chem. Phys. 66, 3045 (1977)]을 적용하여 계산할 수 있다. 밀도범함수 이론의 장점은 전자밀도만 고려하면 되므로 하트리-포크(Hartree-Fock) 방법과 비슷한 수준의 계산량으로 더 정확한 결과를 얻어낼 수 있다. 전자들의 교환-상관에너지를 계산을 위해 교환 범함수들과 상관 범함수들의 조합을 사용하여 계산량을 늘이지 않고도 더 향상된 결과를 얻을 수 있다.

- [0032] 최적의 양자역학 계산방법을 선별하기 위해 선행하여 시도하였던 계산이론은 상기에 언급된 하트리-포크 방법, 다양한 포스트 하트리-포크 방법, 가우시안(G2, G3) 방법, 다양한 범함수 조합의 밀도범함수 이론 등이다. 이 중에서 계산시간 대비 가장 성능이 우수한 밀도범함수 이론의 한가지 방법을 선별하였다. 따라서, 본 발명에서는 상용 양자역학 계산 프로그램을 이용하여 지정된 밀도범함수 이론의 계산방법을 적용하여 분자구조에 대한 최적화 및 진동수 계산을 수행할 수 있고, 이를 통해 상기 구획 내에 포함된 원자의 결합, 각도, 면각도 및 이들 중 2 이상의 조합으로 이루어진 군에서 선택된 하나의, 바람직하게는, 결합, 각도 및 면각도의 조합으로 구성된 입력값을 얻을 수 있다. 마찬가지로, 양자역학 이외의 전산재료 계산방법을 이용한 구조 샘플링을 적용할 수도 있다.
- [0033] 상기 (c) 단계에서는 상기 입력값 중 일부를 이용하여 인공신경망을 학습시켜 새로운 분자구조를 예측, 모사할 수 있는 포텐셜을 생성할 수 있다.
- [0034] 인공신경망 모델은 인간의 뇌의 학습 방법을 모방한 수학적 모델이다. 대뇌의 신경망과 같이 뉴런으로의 입력 신호와 뉴런으로부터의 출력 신호를 각각의 입/출력 노드로 표현하고 시냅스의 결합 세기에 의해 노드 간 연결 강도가 결정된다. 인공신경망에서는 흔히 다층 퍼셉트론(multilayer perceptron)이 많이 적용되며, 이 경우 입력층(input layer), 출력층(output layer) 그리고 하나 이상의 은닉층(hidden layer)으로 구성된다.
- [0035] 각 층은 다수의 뉴런으로 구성되어 있고, 각 층의 뉴런은 연결 가중치로 다른 층의 뉴런과 연결되어 있다. 인공신경망에서의 하나의 뉴런은 생물학적 신경망에서의 하나의 생물학적 신경세포의 역할을 한다.
- [0036] 입력층에 입력된 값과 입력층과 은닉층 사이의 연결 가중치의 곱이 더해져 은닉층 뉴런에 입력되고, 입력된 값은 뉴런 내의 비선형 활성화 함수를 거쳐 입력된 값과는 다른 값을 출력한다. 또한, 출력 결과를 다른 은닉층 혹은 출력층의 다른 뉴런에게 전달하는 과정을 피드포워드(feed forward) 연산이라 하며, 산출된 출력 값과 원하는 타겟 값이 일치하도록 신경망 내의 연결 가중치의 값을 수정하는 과정을 반복적으로 수행하게 되는데 이를 학습(learning)이라고 정의할 수 있다. 여기서 연결 가중치의 값이 갱신은 역전파 학습 알고리즘(back propagation learning algorithm, rumelhart et al, 1986)에 의해 수행될 수 있다. 지정한 에포크(epoch)에 도달하거나 훈련 에러가 목표 값에 도달한 경우 연산은 멈추게 된다.
- [0037] 상기 인공신경망에 대한 입력값은 훈련 집합(training set) 또는 훈련 데이터(training data)라고 하며, 상기 (b) 단계에서 얻어진 입력값 중 일부, 바람직하게는, 10~50%가 훈련 집합으로 선별, 적용될 수 있다. 이러한 선별, 적용에 특별한 기준이 있는 것은 아니지만, 상기 표본 중 상호 인접하여 일정 수준의 중첩이 허용된 구획을 포함하는 하나 이상의 표본에 대한 양자역학 계산 결과가 포함되도록 훈련 집합을 구성함으로써 분자 내 및 분자 간 상호작용을 더 정확하고 신속하게 예측, 모사할 수 있다.
- [0038] 상기 (d) 단계에서는 상기 인공신경망을 이용하여 상기 분자 중 상기 표본을 제외한 부분의 구조를 예측할 수 있고, 상기 (d) 단계에서 상기 분자 중 상기 표본을 제외한 부분에 대해 예측된 구조와 양자역학 또는 전산재료 계산값 간의 오차가 10% 이내일 때까지 상기 (a) 내지 (d) 단계를 반복할 수 있다.
- [0039] 모사방법의 검증
- [0040] 폴리에틸렌 분자를 20개의 원자를 포함하는 구획으로 분할하되, 인접 구획 간 20~25%의 중첩을 허용하였고, 분할된 구획을 약 10~30개씩 임의로 혼합하여 20,000개의 표본을 얻었다.
- [0041] 얻어진 20,000개의 표본 각각에 대한 양자역학 계산을 통해 분자의 구조를 특징지을 수 있는 원자의 결합, 각도, 면각도에 대한 20,000개의 데이터 집합(data set)을 얻었다. 상기 데이터 집합 중 7,500개를 훈련 집합(training set)으로 선정하여 인공신경망을 학습시켜 새로운 분자구조를 예측, 모사할 수 있는 포텐셜을 생성하였다.
- [0042] 학습된 인공신경망을 이용하여 데이터 집합 중 훈련 집합으로 선정되지 않은 100개의 임의의 데이터 집합에 대한 분자구조를 예측하였고, 그 결과를 도 6에 나타내었다. 도 6을 참고하면, 20,000개의 데이터 집합 중 일부인 7,500개를 훈련 집합으로 선정하여 인공신경망을 학습시킨 경우, 학습된 인공신경망이 약 97% 이상의 정확도로 새로운 분자구조를 예측한 것으로 나타나, 본 발명에 따른 분자구조의 모사방법은 1,000개 이상의 원자를 포함하는 고분자, 거대분자의 분자 내 및 분자 간 상호작용을 고려하여 그 구조를 매우 정확하고 신속하게 예측, 모사할 수 있음을 알 수 있다.
- [0043] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해

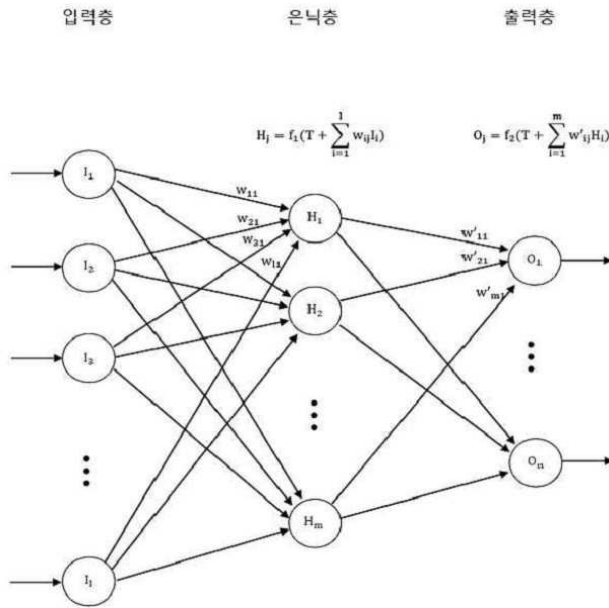
할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

[0044]

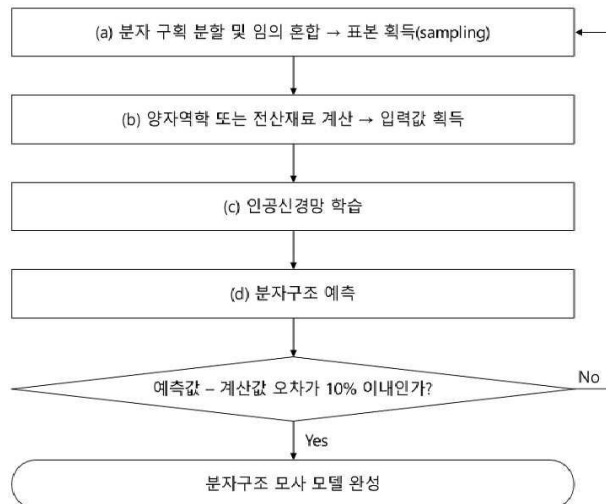
본 발명의 범위는 후술하는 청구범위에 의하여 나타내어지며, 청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

도면

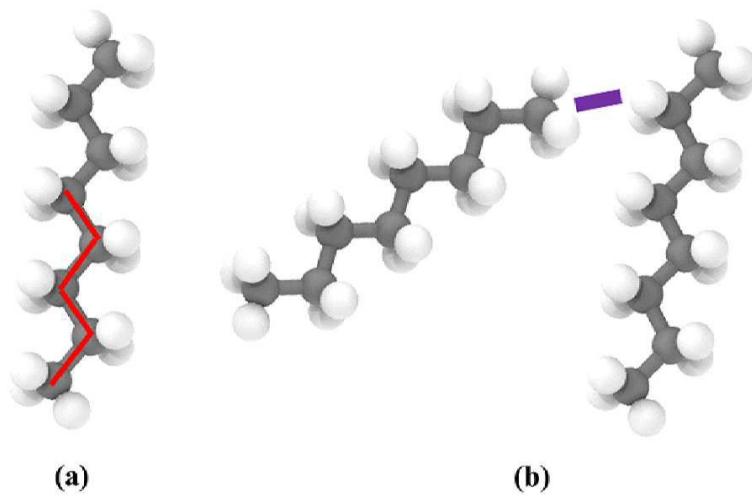
도면1



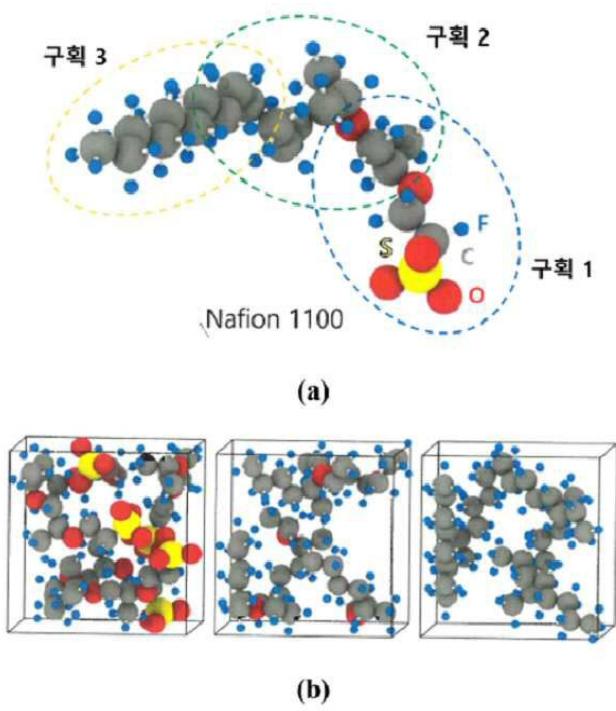
도면2



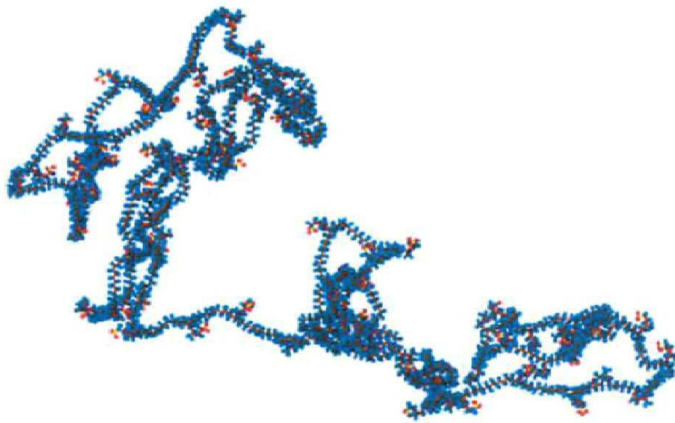
도면3



도면4



도면5



도면6

