



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2022년07월14일  
(11) 등록번호 10-2420872  
(24) 등록일자 2022년07월11일

(51) 국제특허분류(Int. Cl.)  
G16H 20/17 (2018.01) A61M 5/168 (2006.01)  
G06N 20/00 (2019.01) G16H 50/20 (2018.01)  
G16H 50/70 (2018.01)  
(52) CPC특허분류  
G16H 20/17 (2021.08)  
A61M 5/168 (2013.01)  
(21) 출원번호 10-2020-0178132  
(22) 출원일자 2020년12월18일  
심사청구일자 2020년12월18일  
(65) 공개번호 10-2022-0087763  
(43) 공개일자 2022년06월27일  
(56) 선행기술조사문헌  
KR1020110119409 A\*  
KR1020160022874 A\*  
KR1020200063960 A\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
포항공과대학교 산학협력단  
경상북도 포항시 남구 청암로 77 (지곡동)  
(72) 발명자  
박성민  
경상북도 포항시 남구 청암로 77, 창의IT융합공학  
(지곡동)  
이승현  
대구광역시 수성구 수성로24길 63 (상동)  
(74) 대리인  
특허법인태백, 특허법인 다해

전체 청구항 수 : 총 5 항

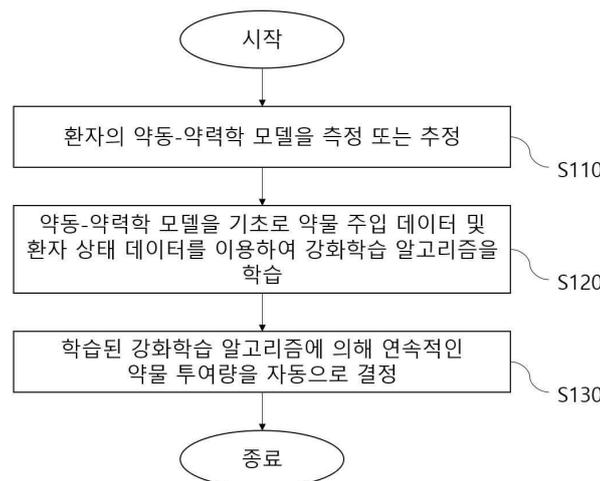
심사관 : 김진률

(54) 발명의 명칭 강화학습과 약동-약리학 모델을 이용한 연속 약물 투여량 결정 방법

(57) 요약

본 발명의 일 실시예에 따른 강화학습과 약동-약리학 모델을 이용한 연속 약물 투여량 결정 방법은, 환자의 약동-약리학 모델을 측정 또는 추정하는 단계; 상기 약동-약리학 모델을 기초로 약물 주입 데이터 및 환자 상태 데이터를 이용하여 강화학습 알고리즘을 학습하는 단계; 및 학습된 강화학습 알고리즘에 의해 연속적인 약물 투여량을 자동으로 결정하는 단계를 포함할 수 있다.

대표도 - 도1



(52) CPC특허분류

*G06N 20/00* (2021.08)

*G16H 50/20* (2018.01)

*G16H 50/70* (2018.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711110873
과제번호	2020R1A2C2005385
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)(R&D)
연구과제명	생체대사 모방형 완전자율 인공체장
기여율	1/1
과제수행기관명	포항공과대학교
연구기간	2020.03.01 ~ 2021.02.28
공지예외적용	: 있음

---

## 명세서

### 청구범위

#### 청구항 1

프로세싱 장치에 의해 수행되는 강화학습과 약동-약리학 모델을 이용한 연속 약물 투여량 결정 방법에 있어서, 환자의 약동-약리학 모델을 측정 또는 추정하는 단계;

상기 약동-약리학 모델을 기초로 약물 주입 데이터 및 환자 상태 데이터를 이용하여 강화학습 알고리즘을 학습하는 단계; 및

학습된 강화학습 알고리즘에 의해 연속적인 약물 투여량을 자동으로 결정하는 단계를 포함하며,

상기 강화학습 알고리즘은 약동-약력 특징을 감가율로 사용하는 것을 특징으로 하는 강화학습과 약동-약리학 모델을 이용한 연속 약물 투여량 결정 방법.

#### 청구항 2

삭제

#### 청구항 3

제 1 항에 있어서,

상기 강화학습 알고리즘은 상기 약동-약리학 모델에서 약효를 단기 약효 및 누적 약효로 구분하며, 단기 약효의 경우 약동-약리학 곡선을 상기 감가율로 사용하고, 누적 약효는 상기 약동-약리학 곡선의 적분값을 상기 감가율로 사용하는 것을 특징으로 하는 강화학습과 약동-약리학 모델을 이용한 연속 약물 투여량 결정 방법.

#### 청구항 4

제 1 항에 있어서,

상기 감가율은 단조 감가율  $r^n$ 과 약동-약리학 모델  $f_n$ 을 곱한 결합 감가율  $r^n f_n$ 인 것을 특징으로 하는 강화학습과 약동-약리학 모델을 이용한 연속 약물 투여량 결정 방법.

#### 청구항 5

제 4 항에 있어서,

상기 강화학습 알고리즘은 시간  $t$ 에서의 환자의 상태  $s_t$ 에서 약물 투여량으로서  $a_t$ 를 선택하는 경우에, 이후 투여된 약물에 의해 시간에 따라 변화하는 환자의 상태에 따라 보상  $R_t, R_{t+1}, \dots, R_{t+n}$ 을 주며, 상기 보상은 상기 결합 감가율에 의해 시간에 따라 감가되며,

상태  $s_t$ 에서의 약물 투여량  $a_t$ 의 평가 및 알고리즘 업데이트는, 상기 보상  $R_t, R_{t+1}, \dots, R_{t+n}$ 에 결합 감가율을 각각 곱해준 다음 모두 더하여 구한  $G_{r,t}$ 로 이루어지는 것을 특징으로 하는 강화학습과 약동-약리학 모델을 이용한 연속 약물 투여량 결정 방법.

#### 청구항 6

제 1 항에 있어서,

상기 감가율은 단조 감가율  $r^n$ 과 누적 약동-약력학 모델  $F_n$ 을 곱한 결합 감가율  $r^n F_n$ 인 것을 특징으로 하는 강화 학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법.

**발명의 설명**

**기술 분야**

[0001] 본 출원은 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법에 관한 것이다.

**배경 기술**

[0002] 약물주입 펌프를 사용한 지속적인 약물의 투여는 암, 당뇨병, 통증 관리, 마취 등 약물을 사용해 환자의 상태를 장기간 조절할 필요가 있는 다양한 의료분야에서 사용되는 관리 및 치료 방법이다.

[0003] 일반적으로 약물주입 펌프의 주입량은 의료진이 상시 관찰을 하며 환자 상태에 따라 직접 입력을 하거나 사전에 정해진 시간별 주입량 프로파일을 수정 없이 사용한다.

[0004] 최근에는 의료인력의 부족 문제의 대안으로 또는 효율성을 위해 자동화된 주입량 결정 및 주입을 하는 폐루프(closed-loop) 알고리즘의 약물주입 펌프 적용이 많이 연구되고 있다.

[0005] 약물에 대한 환자의 약리학적 특성은 환자 개개인에 따라 상이하하며 환자의 상태에 따라서도 변동이 크다는 문제가 있지만, 강화학습(Reinforcement Learning)과 같은 인공지능 학습형 알고리즘이 이를 어느 정도 해결할 수 있다.

[0006] 그러나, 주입한 약물의 약효의 시간 지연 문제는 알고리즘이 급격한 환자상태 변화에 대응하는 것을 어렵게 하며, 이로 인한 약물주입 자동화 알고리즘의 약물 과잉주입 위험이 항상 존재한다.

**선행기술문헌**

**특허문헌**

(특허문헌 0001) 공개특허공보 제10-2011-0119409호(2011.11.02. 공개)

**발명의 내용**

**해결하려는 과제**

[0007] 따라서, 당해 기술분야에서는 자동화된 약물주입 펌프를 통해 연속적으로 약물을 주입하면서도 약물지연을 고려하는 학습형 알고리즘을 구현하기 위한 방안이 요구되고 있다.

**과제의 해결 수단**

[0008] 상기 과제를 해결하기 위해서, 본 발명의 일 실시예는 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법을 제공한다.

[0009] 상기 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법은, 환자의 약동-약력학 모델을 측정 또는 추정하는 단계; 상기 약동-약력학 모델을 기초로 약물 주입 데이터 및 환자 상태 데이터를 이용하여 강화 학습 알고리즘을 학습하는 단계; 및 학습된 강화학습 알고리즘에 의해 연속적인 약물 투여량을 자동으로 결정하는 단계를 포함할 수 있다.

[0011] 덧붙여 상기한 과제의 해결수단은, 본 발명의 특징을 모두 열거한 것이 아니다. 본 발명의 다양한 특징과 그에 따른 장점과 효과는 아래의 구체적인 실시형태를 참조하여 보다 상세하게 이해될 수 있을 것이다.

**발명의 효과**

[0012] 본 발명의 실시예에 따르면, 개별 환자의 연속 약물주입 알고리즘을 학습할 수 있으며, 자동화된 약물주입 알고리즘을 과다투여의 위험 없이 지속적으로 업데이트 할 수 있다.

**도면의 간단한 설명**

[0013] 도 1은 본 발명의 일 실시예에 따른 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법의 흐름도이다.

도 2는 본 발명의 일 실시예에 따라 연속 약물 투여량의 결정에 강화학습 알고리즘을 적용하기 위한 문제 정의 모델을 도시하는 도면이다.

도 3은 본 발명의 일 실시예에 따라 약동-약력학 모델을 사용한 경우 강화학습 알고리즘의 감가율을 설명하기 위한 도면이다.

도 4는 본 발명의 일 실시예에 따라 누적 약동-약력학 모델을 사용한 경우 강화학습 알고리즘의 감가율을 설명하기 위한 도면이다.

도 5는 본 발명의 일 실시예에 따른 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법을 적용하기 전과 후의 효과를 비교하는 도면이다.

**발명을 실시하기 위한 구체적인 내용**

[0014] 이하, 첨부된 도면을 참조하여 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있도록 바람직한 실시예를 상세히 설명한다. 다만, 본 발명의 바람직한 실시예를 상세하게 설명함에 있어, 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략한다. 또한, 유사한 기능 및 작용을 하는 부분에 대해서는 도면 전체에 걸쳐 동일한 부호를 사용한다.

[0015] 덧붙여, 명세서 전체에서, 어떤 부분이 다른 부분과 '연결'되어 있다고 할 때, 이는 '직접적으로 연결'되어 있는 경우뿐만 아니라, 그 중간에 다른 소자를 사이에 두고 '간접적으로 연결'되어 있는 경우도 포함한다. 또한, 어떤 구성요소를 '포함'한다는 것은, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있다는 것을 의미한다.

[0017] 도 1은 본 발명의 일 실시예에 따른 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법의 흐름도이다.

[0018] 도 1을 참조하면, 본 발명의 일 실시예에 따른 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법은, 환자의 약동-약력학 모델을 측정 또는 추정하는 단계(S110), 약동-약력학 모델을 기초로 약물 주입 데이터 및 환자 상태 데이터를 이용하여 강화학습 알고리즘을 학습하는 단계(S120) 및 학습된 강화학습 알고리즘에 의해 연속적인 약물 투여량을 자동으로 결정하는 단계(S130)를 포함하여 구성될 수 있다.

[0019] 여기서, 강화학습 알고리즘은 약리학에서 약효를 의미하는 약동/약력 (Pharmacokinetic/Pharmacodynamic 또는 PK/PD) 특징을 보상의 현재로의 가치 환원에 해당하는 감가율(discount rate)로서 사용할 수 있다. 또한, 약동-약력학 모델을 측정 또는 추정하는 것은 통상의 기술자에게 알려진 기술에 따라 수행될 수 있는 바, 이에 대한 구체적인 설명은 생략한다.

[0021] 또한, 약효를 단기 약효와 장기 약효로 구분하여 적용할 수도 있으며, 이 경우 단기 약효는 PK/PD 곡선을 그대로 감가율로서 사용하고, 장기 약효(즉, 누적 약효)는 PK/PD 곡선의 적분값을 감가율로서 사용할 수 있다.

[0023] 도 1을 참조하여 상술한 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법은 강화학습 알고리즘의 실행이 가능한 프로세싱 장치에 의해 수행될 수 있다.

[0025] 이하, 도 2 내지 도 4를 참조하여 본 발명의 실시예에 따른 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법에 대해 보다 구체적으로 설명한다.

[0027] 도 2는 본 발명의 일 실시예에 따라 연속 약물 투여량의 결정에 강화학습 알고리즘을 적용하기 위한 문제 정의 모델을 도시하는 도면이다.

[0028] 도 2를 참조하면, 강화학습 알고리즘에서 환자의 상태는 정상상태( $S_{normal}$ ), 과소주입 상태( $S_{hyper}$ ) 및 과잉주입 상

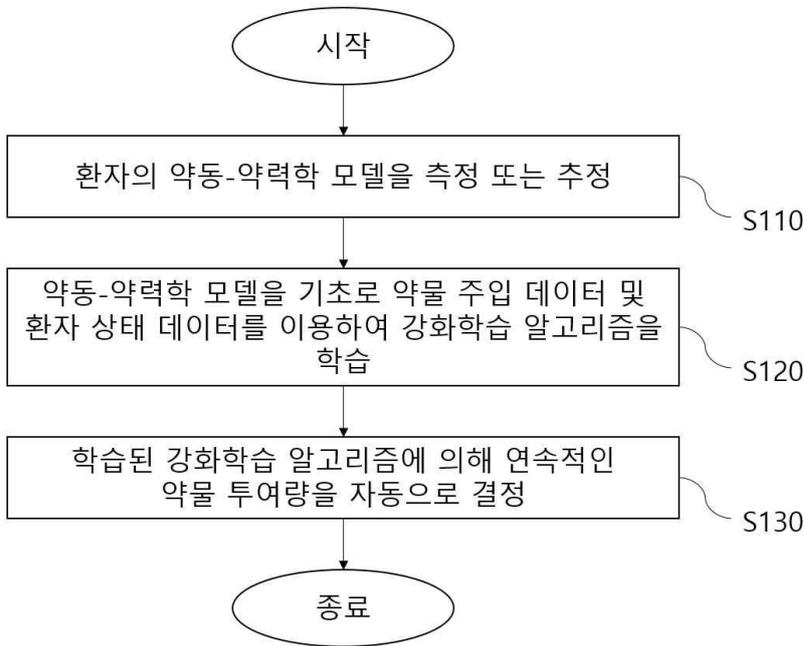
태( $S_{\text{hypo}}$ )를 포함할 수 있다. 여기서, 정상상태( $S_{\text{normal}}$ )는 적절한 양의 약물이 주입된 상태를 의미하고, 과소주입 상태( $S_{\text{hyper}}$ )는 약물의 과소주입으로 인한 상태를 의미하고, 과잉주입 상태( $S_{\text{hypo}}$ )는 약물의 과잉주입으로 인한 상태를 의미한다.

- [0029] 또한, 선택 가능한 약물 주입량(Drug dose)은 0mg, 1mg 및 2mg이라고 할 때, 연속 약물 투여량 결정에 강화학습 알고리즘을 적용하기 위한 문제 정의 모델의 예시는 도 2에 도시된 바와 같다. 여기서, 강화학습은 장기적으로 예상되는 보상이 최대가 되는 결정을 학습하는 인공지능 알고리즘으로, 연속된 결정을 최적화하기 위해 적합하다.
- [0030] 같은 양의 약물을 투여하더라도 항상 동일한 약효가 나타나는 것이 아니므로 환자의 상태는 상태 전이 확률(state transition probability)에 따라 변하며, 이렇게 상태가 변할 때 변하는 상태에 따라 상이한 보상(reward)이 주어질 수 있다.
- [0031] 예를 들어, 정상상태( $S_{\text{normal}}$ )에서 2mg의 약물을 선택할 경우 90%의 확률로 과잉주입 상태( $S_{\text{hypo}}$ )로 전이되며, 이 경우 -2의 보상을 받게 된다.
- [0032] 반면, 정상상태( $S_{\text{normal}}$ )에서 0mg을 선택할 경우 90%의 확률로 과소주입 상태( $S_{\text{hyper}}$ )로 전이되며, 이 경우 -1의 보상을 받게 된다.
- [0033] 그러나, 정상상태( $S_{\text{normal}}$ )에서 1mg의 약물을 선택할 경우 100%의 확률로 정상상태( $S_{\text{normal}}$ )를 유지할 수 있고, 이 경우 1의 보상을 받게 된다.
- [0034] 이처럼, 환자의 치료 기록과 상태에 따른 보상 기준만 주면, 강화학습 알고리즘은 반복적인 업데이트를 통해 정상상태( $S_{\text{normal}}$ )에서는 1mg을, 과잉주입 상태( $S_{\text{hypo}}$ )에서는 0mg을, 그리고 과소주입 상태( $S_{\text{hyper}}$ )에서는 2mg을 주입해야 하는 것을 학습할 수 있다.
- [0035] 이와 같은 원리는 훨씬 복잡한 실제 환자의 상태, 약물의 종류, 약물의 주입량, 그리고 약물에 대한 반응에도 동일하게 적용될 수 있다.
- [0036] 그러나, 이러한 원리만으로는 약효 지연에 대한 대응이 어렵다는 한계가 있다.
- [0038] 약동학(pharmacokinetics)은 약물의 흡수, 분포, 생체 내 변화 및 배설을 연구하는 분야이고, 약력학(pharmacodynamics)은 생체에 대한 약물의 생리학적 및 생화학적 작용과 그 작용기전, 즉 약물이 일으키는 생체의 반응을 주로 연구하는 분야이다. 즉, 약동학은 주입한 약물의 시간에 따른 혈중 농도 변화에 해당하고, 약력학은 혈중 약물 농도에 따른 약효의 변화에 해당하며, 이를 합쳐서 시간에 따른 약효와 변화를 약동-약력학(PK-PD) 모델이라 한다.
- [0039] 한편, 일반적인 강화학습 알고리즘은 현재의 행동 이후 미래에 어떤 보상을 받았으며, 이 보상들이 현재의 행동에 얼마나 영향을 받았는가를 고려하여 평가 및 업데이트된다.
- [0040] 그러나, 연속 행동 결정 모델에서는 현재 행동의 결과는 일반적으로 시간이 지날수록 그 자체의 영향력이 줄어들며, 행동 이외의 외부 요소에 의해서도 희석된다.
- [0041] 따라서, 행동 결정 이후에 받는 미래의 보상들에 대해서 더 나중에 받은 보상은 더 많이 감가하여 고려하며, 이는 0에서 1 사이의 감가율  $r$ 을 보상에 거듭제곱하여 이루어진다. 즉, 행동으로부터  $n$ 단계 이후에 받는 보상  $R$ 은  $r^n$ 만큼 감가된  $r^n \times R$ 로서 알고리즘의 업데이트에 반영된다.
- [0043] 본 발명에서는 이러한 행동에 따른 보상의 시간에 대한 감가의 개념을 투여한 약물의 시간에 따른 약효 변화의 개념인 약동-약력학과 결합하여 적용하고자 한다.
- [0045] 도 3은 본 발명의 일 실시예에 따라 약동-약력학 모델을 사용한 경우 강화학습 알고리즘의 감가율을 설명하기 위한 도면이다.
- [0046] 도 3을 참조하면, 시간  $t$ 에서의 환자의 상태  $s_t$ 에서 알고리즘이 약물 투여량으로서  $a_t$ 를 선택하는 경우에, 이후 투여된 약물에 의해 환자의 상태는 시간에 따라 변화하며 이에 따른 보상  $R_t, R_{t+1}, \dots, R_{t+n}$ 이 주어지는 경우를 나타낸다.

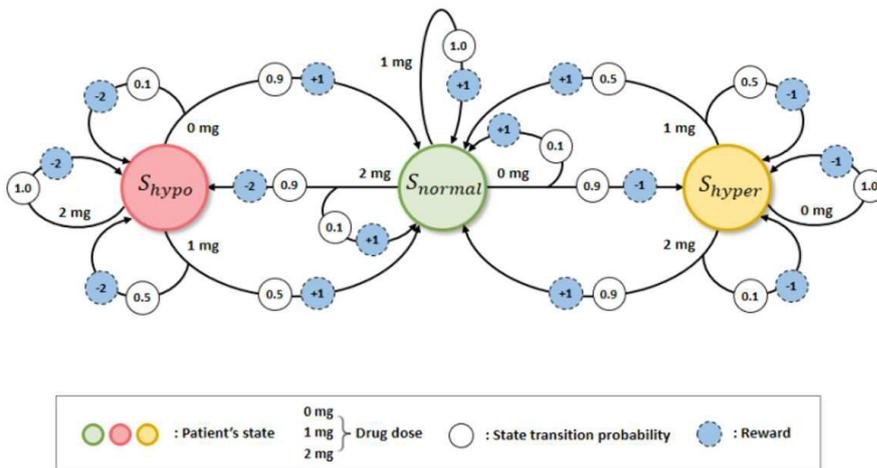
- [0047] 적절한 약물 투여량을 학습하는 과정으로서 투여량  $a_t$ 의 적절성을 평가하기 위해, 이 보상들을 행동의 영향력을 고려하여 시간  $t$ 로 감가시키다.
- [0048] 회색 실선은 외부 요소에 의한 시간에 따른 행동 영향력의 회색을 의미하는 것으로, 일반적으로 강화학습에서 사용되는 단조 감가율  $r^n$ 을 나타낸다.
- [0049] 또한, 붉은 점선은 약동-약력학 모델  $f_n$ 으로, 일반적으로 약물주입 후 상승하다가 정점 후 감소하는 양상을 보여 준다.
- [0050] 마지막으로, 붉은 실선은 본 발명에서 제안하는 연속 약물 투여를 위한 결합 감가율로, 상술한 단조 감가율  $r^n$ 과 약동-약력학 모델  $f_n$ 을 곱하여  $r^n f_n$ 으로 나타낼 수 있다.
- [0051] 따라서, 상태  $s_t$ 에서의 약물투여량  $a_t$ 의 평가 및 알고리즘 업데이트는, 시간에 따른 이후 보상  $R_t, R_{t+1}, \dots, R_{t+n}$ 에 결합 감가율을 각각 곱해준 다음 모두 더하여 구한  $G_{t,1}$ 로 이루어질 수 있다.
- [0053] 한편, 시간에 따른 단일 시점의 약효를 평가하는 것이 아니라, 누적 약효를 평가하기 위해서도 상술한 개념을 적용할 수 있다.
- [0055] 도 4는 본 발명의 일 실시예에 따라 누적 약동-약력학 모델을 사용한 경우 강화학습 알고리즘의 감가율을 설명하기 위한 도면으로, 도 4에 도시된 누적 약동-약력학 모델(cumulative PK-PD model)을 사용하여 도 3에 도시된 바와 동일하게 결합 감가율을 산출하여 사용할 수 있다.
- [0057] 도 5는 본 발명의 일 실시예에 따른 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법을 적용하기 전과 후의 효과를 비교하는 도면으로, 미국 FDA에서 동물실험 대체를 승인한 가상의 당뇨병 환자 시뮬레이터에서 강화학습을 사용하여 연속 인슐린 주입속도를 결정하는 알고리즘을 학습한 결과를 나타낸다.
- [0058] 또한, 도 5의 (a)는 본 발명에 따른 결합 감가율을 적용하기 전의 결과를 나타내고, (b)는 본 발명에 따른 결합 감가율을 적용한 후의 결과를 나타내는 것으로, (a) 및 (b)에 도시된 4개의 그래프는 위에서부터 순차로 각각 시간에 따른 환자의 혈당값(BG), 식사(CHO), 약물주입 펌프로 주입한 인슐린양(Ins), 그리고 강화학습 알고리즘이 추정한 체내 잔여 인슐린양(IOB)을 나타내는 것이다.
- [0059] 도 5를 참조하면, 강화학습 알고리즘은 환자의 혈당값을 기초로 인슐린 주입 속도를 결정하여 약물주입 펌프로 인슐린을 주입하게 된다.
- [0060] 또한, 본 발명의 실시예에 따른 연속 약물 투여량 결정 방법이 과잉주입의 위험 없이 자동화된 약물 투여를 가능하게 하는지 평가하기 위해, 아침/점심/저녁 3끼의 식사를 2일 차부터 가상환자에 제공하였으며 식사량은 알고리즘에 입력되지 않았다.
- [0061] 도 5의 (a) 및 (b)를 비교하면, 점심에 과다한 식사를 하여 혈당이 빠르게 상승하는 경우, 결합 감가율 적용 전에는 인슐린을 과잉 주입하여 17시~19시 사이에 저혈당(가장 위의 그래프에서 아래쪽 빨간색 영역)에 빠지게 된다. 이에 반해, 결합 감가율 적용 후에는, 점심 식사 후에 빠르게 인슐린을 주입하되 필요한 만큼의 적정량만 주입 후 투여량을 다시 줄여주므로 저혈당이 오지 않고 혈당을 정상 범위(가장 위의 그래프에서 초록색 영역)에서 잘 유지함을 확인할 수 있다.
- [0063] 상술한 바와 같은 본 발명의 실시예에 따른 강화학습과 약동-약력학 모델을 이용한 연속 약물 투여량 결정 방법은 정밀의료의 일환으로서 약물주입 펌프의 약물투여 개인화에 활용될 수 있다.
- [0064] 또한, 약물 약효지연을 고려하여 과잉투여의 위험이 없는 자동화된 약물투여가 가능하므로 만성질환 환자의 질병관리를 위한 원격의료에도 활용될 수 있다.
- [0065] 대표적인 예로서, 암, 당뇨병, 통증 관리, 마취 등의 의료 현장에 사용 가능하며, 특히 당뇨병의 경우 식사량의 입력이 필요 없는 완전 자율화된 인공체장의 구현을 위해 적용될 수 있다.
- [0067] 본 발명은 전술한 실시예 및 첨부된 도면에 의해 한정되는 것이 아니다. 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 있어, 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 본 발명에 따른 구성요소를 치환, 변형 및 변경할 수 있다는 것이 명백할 것이다.

도면

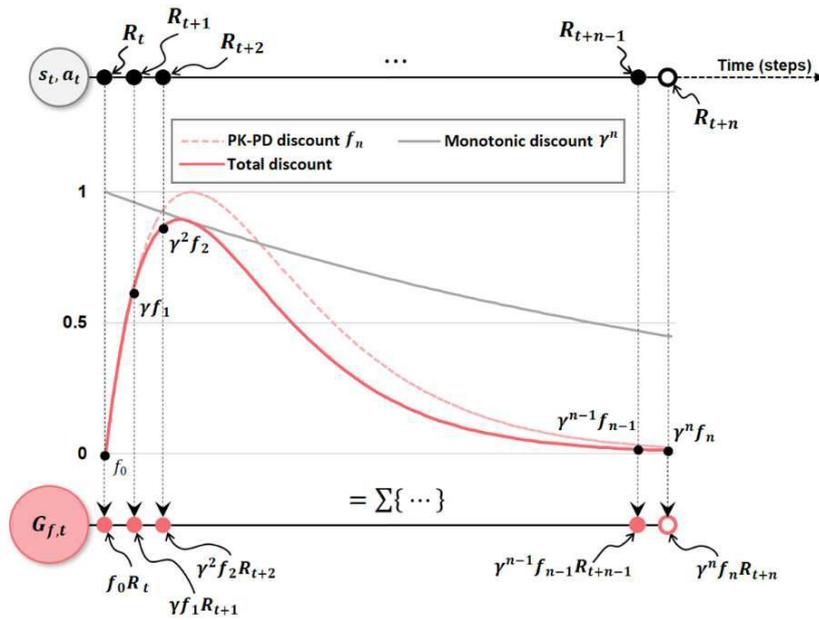
도면1



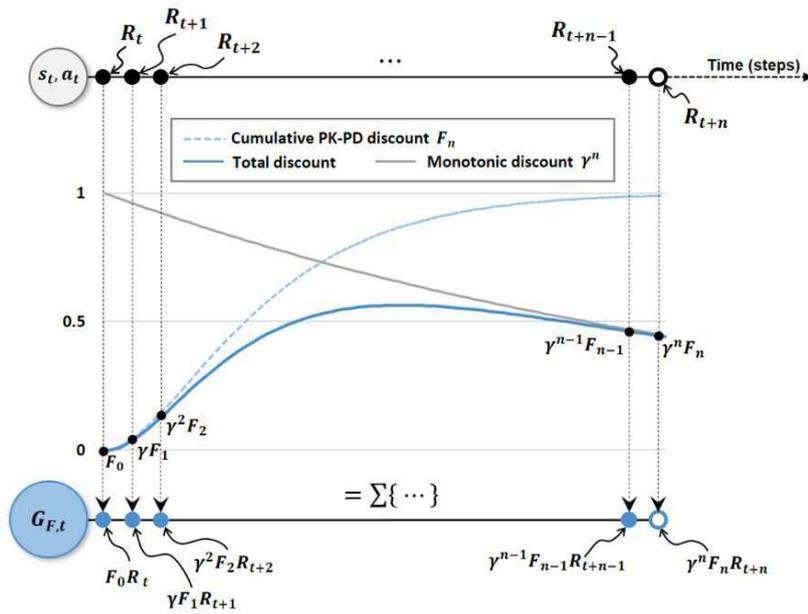
도면2



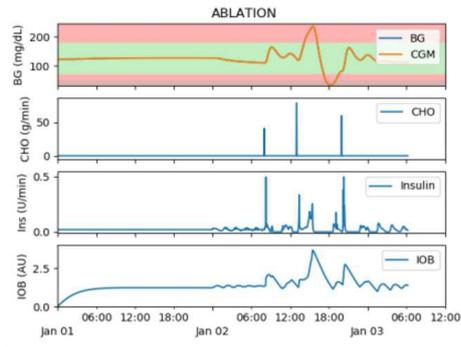
도면3



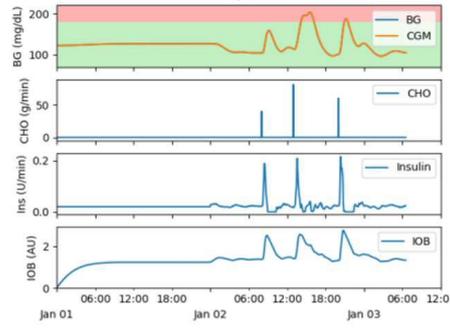
도면4



도면5



(a)



(b)