



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2022년06월13일

(11) 등록번호 10-2408011

(24) 등록일자 2022년06월08일

(51) 국제특허분류(Int. Cl.)

B25J 9/10 (2006.01) B25J 13/06 (2006.01)

B25J 19/02 (2006.01) B25J 9/16 (2006.01)

(52) CPC특허분류

B25J 9/1075 (2013.01)

B25J 13/065 (2013.01)

(21) 출원번호 10-2020-0182336

(22) 출원일자 2020년12월23일

심사청구일자 2020년12월23일

(56) 선행기술조사문헌

JP2012125847 A*

W02020055342 A1*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

울산대학교 산학협력단

울산광역시 남구 대학로 93(무거동)

(72) 발명자

안경관

울산광역시 남구 대공원로 207, 102동 402호(신정동, 대공원 월드메르디앙)

보, 풍 팻

울산광역시 남구 대학로33번길 30-29, 102호(무거동)

(74) 대리인

김정수

전체 청구항 수 : 총 11 항

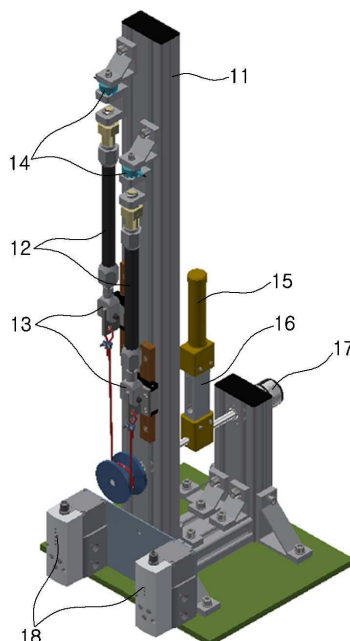
심사관 : 이상용

(54) 발명의 명칭 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법

(57) 요약

본 발명은 공압 인공근육(Pneumatic Artificial Muscle ; PAM) 시스템의 인인간-로봇 협력(human-robot cooperative) 제어장치 및 방법에 관한 것으로, 본 발명에 따르면, 인간근육의 움직임에 최대한 가깝게 인공근육의 움직임을 구현하는 데에만 중점을 두고 있는 것이 대부분이었던 종래기술의 공압 인공근육(PAM) 시스템 및 이

(뒷면에 계속)

대표도 - 도110

를 이용한 로봇들의 한계를 해결하기 위해, 인간과 로봇의 상호작용을 측정하여 제어스틱의 움직임에 대응하는 힘을 인공근육에 전달하고 인공근육의 움직임에 대응하는 힘을 제어스틱에 전달하는 동시에, 강화학습(reinforcement learning)을 이용한 학습을 통해 각 개인에게 적합한 최적의 모델을 결정하는 처리가 수행되도록 구성됨으로써, 공압 인공근육(PAM) 모듈에 대하여 물리적 인간-로봇 상호작용(physical human-robot interaction ; pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있는 동시에, 다양한 사용자에게 적용할 수 있도록 구성되는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법이 제공된다.

(52) CPC특허분류

B25J 19/02 (2013.01)

B25J 9/104 (2013.01)

B25J 9/161 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711110290
과제번호	NRF-2020R1A2B5B03001480
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	유체마찰전기 기반의 고효율 에너지 수확장치 개발
기 여 율	1/1
과제수행기관명	울산대학교
연구기간	2020.03.01 ~ 2023.02.28

명세서

청구범위

청구항 1

강화학습(reinforcement learning)을 이용한 공압 인공근육(Pneumatic Artificial Muscle ; PAM) 모듈의 인간-로봇 협력(human-robot cooperative) 제어장치에 있어서,

한 쌍의 공압 인공근육(PAM)이 케이블을 통해 풀리(pully)에 연결되는 것에 의해 서로 대립구성을 이루도록 형성되는 인공근육모듈;

상기 인공근육 모듈의 움직임을 제어하기 위한 제어스틱(control stick);

인공지능(AI)을 이용하여, 강화학습을 이용한 학습을 통해 상기 제어스틱에 가해지는 힘과 상기 제어스틱의 움직임에 대응하여 상기 인공근육모듈의 움직임을 제어하는 동시에, 상기 인공근육모듈의 움직임에 대응하는 힘을 상기 제어스틱에 전달하는 것에 의해 인간-로봇 상호작용(human-robot interaction)을 구현하는 처리가 수행되도록 이루어지는 제어부;

상기 제어장치의 본체를 형성하는 메인프레임;

상기 인공근육모듈의 각각의 공압 인공근육(PAM)을 구동하기 위해 상기 메인프레임에 설치되는 한 쌍의 PAM 액츄에이터(actuator);

상기 인공근육모듈의 각각의 공압 인공근육(PAM)의 움직임을 검출하기 위해 상기 메인프레임에 설치되는 한 쌍의 로드셀(loadcell);

상기 메인프레임에 회전 가능하게 설치되고 일단에 상기 풀리가 연결되는 샤프트;

상기 제어스틱의 움직임을 측정하기 위해 상기 제어스틱의 핸들과 상기 샤프트 사이에 설치되는 인터랙티브 로드셀(interactive loadcell);

상기 제어스틱의 움직임을 측정하기 위해 상기 샤프트의 타단에 설치되는 광학 로터리 엔코더(optical rotary encoder); 및

내부에 구비된 압력센서의 측정값에 근거하여 상기 제어스틱의 움직임에 따른 압력으로 압축공기를 상기 인공근육모듈의 각각의 공압 인공근육(PAM)에 전달하기 위해 상기 메인프레임에 설치되는 한 쌍의 비례 압력조절 밸브(proportional pressure regulator valve)를 포함하여 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

청구항 2

삭제

청구항 3

제 1항에 있어서,

상기 제어부는,

이하의 수학적식을 이용하여, 상기 제어스틱과 상기 인공근육모듈 사이의 인간-로봇 상호작용을 임피던스 모델로 나타내고,

$$\theta_m = \frac{K_h \tau_h}{M_d s^2 + C_d s + K_d}$$

(여기서, Θ_m 은 임피던스 모델의 출력위치(output position), M_d 는 관성(inertia) 파라미터 행렬, C_d 는 댐핑(damping) 파라미터 행렬, K_d 는 강성(stiffness) 파라미터 행렬, s 는 일반 비선형 방해(general nonlinear disturbance), K_h 는 인간 입력힘 이득(human input force gain), τ_h 는 인간 입력힘 이득 K_h 를 가지는 인간 제어노력(human control effort)을 각각 의미함.)

상기 인공근육 모듈이 상기 임피던스 모델에 따라 동작하도록 토크 제어를 수행하는 내부루프(inner loop) 처리와, 상기 임피던스 모델에 대한 최적의 파라미터 값을 산출하는 외부루프(outer loop) 처리가 수행되도록 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

청구항 4

제 3항에 있어서,

상기 제어부는,

상기 외부루프(outer loop) 처리를 통해 산출된 상기 임피던스 모델에 대한 파라미터 값을 보상(reward)으로 적용하여 강화학습을 통해 최적의 파라미터 값을 업데이트 하는 처리가 수행되도록 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

청구항 5

제 4항에 있어서,

상기 내부루프 처리는,

이하의 수학적식을 이용하여, 상기 인공근육 모듈이 상기 임피던스 모델에 따라 동작하도록 하기 위한 제어토크(control torque)(u_c)를 산출하는 처리가 수행되도록 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

$$u_c = \hat{\psi}(q) + K_s \sigma_h - u_r(t)$$

$$\begin{cases} \hat{\psi}(q) = \hat{W}_1^T \phi(\hat{W}_2^T q) \\ \sigma_h = \dot{e}_h + \mu_1 e_h + \mu_2 \int_0^t e_h(\tau) d\tau \end{cases}$$

(여기서, $u_r(t)$ 는 테일러 급수(Taylor series)의 고차항(high-order term)에 대한 견고성(robustness)을 제공하도록 특정된 강건화 신호(robustifying signal)이고, K_s 는 제어이득(control gain)이며, σ_h 는 슬라이딩모드 오

차(sliding mode error)이고, $\hat{\psi}(q)$ 는 $q = [\theta, \dot{\theta}, \ddot{\theta}_m, \ddot{\theta}_m, e_h, \dot{e}_h, \xi]^T$ 및 $\xi = \int_0^t e_h(\tau) d\tau$ 를 가지는 신경망(Neural Network ; NN)을 의미함)

청구항 6

제 5항에 있어서,

상기 외부루프 처리는,

관절각 오차(joint angle error) 및 인간-로봇 접촉력에 근거하여 상태공간(state space)을 형성하고 미리 정해진 간격으로 분할하는 처리가 수행되는 상태공간 및 이산화 처리단계;

상태에 따른 행동집합(act in set)을 정의하는 처리가 수행되는 행동집합 처리단계;

강화학습 알고리즘의 정책을 결정하는 처리가 수행되는 학습 알고리즘 처리단계;

상기 강화학습 알고리즘의 보상을 제공하고 상기 임피던스 모델의 파라미터를 업데이트 하기 위한 증분공식(incremental formular)을 도출하는 처리가 수행되는 업데이트 처리단계; 및

상기 증분공식에 근거하여 상기 임피던스 모델의 파라미터를 업데이트 하는 처리가 수행되는 파라미터 조정단계를 포함하여 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

청구항 7

제 6항에 있어서,

상기 상태공간 및 이산화 처리단계는,

상기 광학 로터리 엔코더 및 상기 인터랙티브 로드셀을 통해 관절각 오차(joint angle error) 및 인간-로봇 접촉력을 각각 측정하고 상태공간(state space)을 형성하며,

경험적 분할법(empirical segmentation method)을 이용하여 상기 관절각 오차와 상기 인간-로봇 접촉력의 연속 값(continuous value)을 M 및 N 간격(interval)으로 각각 나누어 상기 상태공간에 M×N 그리드를 형성하는 처리가 수행되도록 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

청구항 8

제 7항에 있어서,

상기 행동집합 처리단계는,

좌, 우 움직임(movement)에 따라 [1, 0], [0, 1]로 표현되는 두 그룹의 행동으로 상기 행동집합(act in set)을 정의하는 처리가 수행되도록 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

청구항 9

제 8항에 있어서,

상기 학습 알고리즘 처리단계는,

π 와 P를 각각 정책(policy) 및 상태전환 확률(state transition probability)이라 할 때, 이하의 수학적식을 이용하여 ρ -그리디(greedy) 정책을 강화학습 정책으로 적용하는 처리가 수행되도록 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

$$\pi^{\rho} = \begin{cases} \arg \max_a Q(\theta, a) & \text{with probability } 1 - \rho \\ \frac{1}{A(\theta)} & \text{with probability } \rho \end{cases}$$

(여기서, $Q(\theta, a)$ 는 상태 행동-값 함수(state action-value function) 이고, $A(\theta)$ 는 이산 행동집합(discrete action set)에 포함된 총 행동 수를 각각 의미함)

청구항 10

제 9항에 있어서,

상기 업데이트 처리단계는,

이하의 수학적식을 이용하여, 상태 행동-값 함수를 업데이트 하기 위한 증분공식(incremental formular)을 도출하는 처리가 수행되도록 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

$$Q_{t+1}(\theta_t, a_t) = Q_t(\theta_t, a_t) + \alpha (R_{t+1} + \gamma Q_t(\theta_{t+1}, a_{t+1}) - Q_t(\theta_t, a_t))$$

(여기서, t는 이산시간 단계(discrete-time step), α는 단계크기(step size) 파라미터, γ는 할인율(discount-rate) 파라미터, R_{t+1}은 행동 a_t를 취한 후 상태 θ_t에서 상태 θ_{t+1}로 전이(transition)를 가지는 시간 t+1의 보상을 각각 의미함)

청구항 11

제 10항에 있어서,

상기 업데이트 처리단계에서,

상기 보상(R_{t+1})은, 이하의 수학적식으로 나타내지는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치.

$$R_{t+1} = r_{t+1} \mid a_t, \theta_t \rightarrow \theta_{t+1}$$

(여기서, r_{t+1}은 상태-행동쌍(state-action pair) (θ_t, a_t)의 평가(evaluation)에 해당하는 즉시보상(immediate reward)을 의미함)

청구항 12

강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어방법에 있어서,

청구항 1항, 청구항 3항 내지 청구항 11항 중 어느 한 항에 기재된 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치를 이용하여 인공근육 모듈의 제어장치를 구현하는 단계; 및

상기 제어장치를 이용하여 상기 인공근육 모듈의 제어를 수행하는 단계를 포함하여 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어방법.

발명의 설명

기술 분야

[0001] 본 발명은 공압 인공근육(Pneumatic Artificial Muscle ; PAM) 시스템의 인간-로봇 협력(Human-Robot Cooperative) 제어장치 및 방법에 관한 것으로, 더 상세하게는, 인간근육의 움직임에 최대한 가깝게 인공근육의 움직임을 구현하는 데에만 중점을 두고 있는 것이 대부분이었던 종래기술의 공압 인공근육(PAM) 시스템 및 이를 이용한 로봇들의 한계를 해결하기 위해, 공압 인공근육(PAM) 모듈에 대하여 물리적 인간-로봇 상호작용(physical human-robot interaction ; pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있도록 구성되는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법에 관한 것이다.

[0003] 또한, 본 발명은, 상기한 바와 같이 물리적 인간-로봇 상호작용(pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있도록 하기 위해, 케이블을 통해 폴리에 대립적 구성으로 연결된 한 쌍의 공압 인

공근육(PAM)과, 인공근육의 구동을 위한 액츄에이터 및 압축공기를 공급하는 비례압력조절 밸브(proportional pressure regulator valve)와, 인공근육의 움직임을 제어하기 위한 제어스틱(control stick)과, 각각의 공압 인공근육의 움직임을 측정하기 위한 로드셀과, 제어스틱의 움직임을 측정하기 위한 로드셀 및 광학 로터리 엔코더(optical rotary encoder)를 포함하여 구성되고, 미리 설정된 임피던스 모델에 따라 인간과 로봇의 상호작용을 측정하여 제어스틱의 움직임에 대응하는 힘을 인공근육에 전달하고 인공근육의 움직임에 대응하는 힘을 제어스틱에 전달하는 동시에, 강화학습(reinforcement learning)을 이용한 학습을 통해 각 개인에게 적합한 최적의 모델을 결정하는 처리가 수행되도록 구성됨으로써, 로봇이 규정된 동적모델에 따라 동작하도록 하여 상호작용을 단순화할 수 있는 동시에, 다양한 사용자에게 적용할 수 있도록 구성되는 강화학습을 이용한 공압 인공근육 시스템의 인간-로봇 협력 제어장치 및 방법에 관한 것이다.

배경 기술

- [0005] 종래, 일반적으로, 로봇은 각 구동부에 전기모터를 사용하는 것이 대부분이었으나, 최근에는, 인간근육과 같이 외부환경에 맞게 강성을 조절할 수 있는 공압형 인공근육(Pneumatic Artificial Muscle ; PAM)을 적용한 로봇이 많이 이용되고 있다.
- [0007] 이러한 공압형 인공근육(PAM)을 적용한 로봇은 특히 외부환경과의 안정한 접촉이 요구되는 응용분야에 다양하게 적용되고 있으며, 이를 위해, 최근에는, 인공근육의 정확한 모델링에서 근육의 운동제어에 이르기까지 인공근육 및 이를 이용한 로봇에 대한 연구가 활발하게 이루어지고 있다.
- [0009] 여기서, 상기한 바와 같은 인공근육 및 이를 이용한 로봇에 관한 종래기술의 예로는, 예를 들면, 한국 공개특허 공보 제10-2020-0096010호에 제시된 바와 같은 "생체모방형 인공근육모듈과 이를 포함하는 인공근육 어셈블리 및 그 제어방법"이 있다.
- [0011] 더 상세하게는, 상기한 한국 공개특허공보 제10-2020-0096010호는, 길이방향으로 수축운동 또는 이완 운동하는 작용부와, 작용부의 일단에 연결되어 외력의 인가에 따라 탄성 거동하는 탄성부와, 작용부의 타단에 연결되어 작용부의 수축운동또는 이완운동을 구동하는 구동부와, 작용부의 길이가 변화하는 것을 선택적으로 차단하는 잠금부 및 탄성부 및 작용부의 상태를 각각 측정하는 제 1 및 제 2 센서부를 포함하여, 인간의 근육-힘줄 구성단위(Muscle-tendon unit)의 구조를 모사하여 신경학적 제어특성에 근거하여 근수축의 역학적인 거동특성을 실질적으로 재현할 수 있도록 구성되는 생체모방형 인공근육모듈과 이를 포함하는 인공근육 어셈블리 및 그 제어방법에 관한 것이다.
- [0013] 또한, 상기한 바와 같은 인공근육 및 이를 이용한 로봇에 관한 종래기술의 다른 예로는, 예를 들면, 한국 등록특허공보 제10-1834970호에 제시된 바와 같은 "인공근육모듈의 구동장치와 인공근육모듈의 구동방법"이 있다.
- [0015] 더 상세하게는, 상기한 한국 등록특허공보 제10-1834970호는, 고온수탱크와 저온수탱크를 포함하는 유체탱크유닛과, 인공근육모듈과 유체탱크유닛을 연결하는 유체공급라인 및 유체회수라인과, 인공근육모듈과 유체탱크유닛 사이에서 유체를 순환시키는 순환펌프유닛과, 고온수탱크에서 배출되는 유체와 저온수탱크에서 배출되는 유체를 이용하여 인공근육모듈의 형상을 변형시키기 위한 기설정된 작동온도를 갖는 유체를 형성하는 온도조절유닛과, 기설정된 분배조건에 따라 유체회수라인의 유체를 고온수탱크 또는 저온수탱크로 분배하는 유체분배유닛을 포함하여, 인공근육모듈에 충전되는 유체의 온도에 따라 인공근육모듈의 유연한 움직임을 구현하면서 인공근육모듈의 응답성을 빠르게 하고, 종래에 비해 에너지 효율을 향상시킬 수 있도록 구성되는 인공근육모듈의 구동장치와 인공근육모듈의 구동방법에 관한 것이다.
- [0017] 상기한 바와 같이, 종래, 인공근육 및 이를 이용한 로봇에 대하여 여러 가지 기술내용들이 제시된 바 있으나, 상기한 바와 같은 종래기술의 내용들은 다음과 같은 문제점이 있는 것이었다.
- [0019] 더 상세하게는, 최근에는, 상기한 바와 같이 공압형 인공근육(PAM)을 이용한 협업로봇(collaborative robot)의 등장하여 다양한 분야에 적용되고 있으나, 인간과 협업작업을 수행하는 협업로봇의 경우 로봇기구의 운동으로 발생할 수 있는 인간과의 충돌 안전성을 고려하여 로봇의 형태가 설계되어야 한다.
- [0021] 이에 따라, 안전하고 신뢰할 수 있으며 효율적인 물리적 인간-로봇 상호작용(physical human-robot interaction ; pHRI)에 대한 필요성이 증가하고 있으며, 이러한 고성능 pHRI에는 고도의 비선형 로봇(highly nonlinear robot)에 적합한 견고하고(robust) 안정적인(stable) 컨트롤러가 요구된다.
- [0023] 그러나 상기한 바와 같은 종래기술의 내용들은 인간근육의 움직임을 최대한 모방하여 인공근육의 자체의 움직임을 구현하는 데에 중점을 두고 있는 것이 대부분으로, 공압 인공근육 시스템에서 물리적 인간-로봇 상호작용

(pHRI)을 통하여 인공근육의 움직임보다 정확하고 효율적으로 제어할 수 있도록 구성되는 공압 인공근육 시스템의 인간-로봇 협력 제어장치나 제어방법은 제시된 바 없었다.

- [0025] 즉, 인공근육 시스템을 보다 정확하고 효율적으로 제어하기 위하여는, 예를 들면, 인공근육의 움직임량에 따른 힘을 제어장치에 인가하여 움직임이 클수록 저항이 커지게 하거나, 또는, 제어장치에 가해지는 힘이나 이동변위가 클수록 이에 비례하는 힘으로 인공근육의 움직임이 제어되도록 하는 것에 의해, 인공근육을 제어하기 위한 제어장치의 움직임과 인공근육의 움직임이 서로 연동하여 반응하도록 함으로써 보다 정확한 제어가 이루어질 수 있도록 구성되는 새로운 구조의 공압 인공근육 인간-로봇 협력 제어장치 및 방법을 제시하는 것이 바람직하나, 아직까지 그러한 요구를 모두 만족시키는 장치나 방법은 제시되지 못하고 있는 실정이다.

선행기술문헌

특허문헌

- [0027] (특허문헌 0001) 한국 공개특허공보 제10-2020-0096010호 (2020.08.11.)
(특허문헌 0002) 한국 등록특허공보 제10-1834970호 (2018.02.27.)

발명의 내용

해결하려는 과제

- [0028] 본 발명은 상기한 바와 같은 종래기술의 문제점을 해결하고자 하는 것으로, 따라서 본 발명의 목적은, 인간근육의 움직임에 최대한 가깝게 인공근육의 움직임을 구현하는 데에만 중점을 두고 있는 것이 대부분이었던 종래기술의 공압 인공근육(PAM) 시스템 및 이를 이용한 로봇들의 한계를 해결하기 위해, 공압 인공근육(PAM) 모듈에 대하여 물리적 인간-로봇 상호작용(pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있도록 구성되는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법을 제공하고자 하는 것이다.

- [0030] 또한, 본 발명의 다른 목적은, 상기한 바와 같이 물리적 인간-로봇 상호작용(pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있도록 하기 위해, 케이블을 통해 폴리에 대립적 구성으로 연결된 한 쌍의 공압 인공근육(PAM)과, 인공근육의 구동을 위한 액츄에이터 및 압축공기를 공급하는 비례압력조절 밸브(proportional pressure regulator valve)와, 인공근육의 움직임을 제어하기 위한 제어스틱(control stick)과, 각각의 공압 인공근육의 움직임을 측정하기 위한 로드셀과, 제어스틱의 움직임을 측정하기 위한 로드셀 및 광학로터리 엔코더(optical rotary encoder)를 포함하여 구성되고, 미리 설정된 임피던스 모델에 따라 인간과 로봇의 상호작용을 측정하여 제어스틱의 움직임에 대응하는 힘을 인공근육에 전달하고 인공근육의 움직임에 대응하는 힘을 제어스틱에 전달하는 동시에, 강화학습(reinforcement learning)을 이용한 학습을 통해 각 개인에게 적합한 최적의 모델을 결정하는 처리가 수행되도록 구성됨으로써, 로봇이 규정된 동적모델에 따라 동작하도록 하여 상호작용을 단순화할 수 있는 동시에, 다양한 사용자에게 적용할 수 있도록 구성되는 강화학습을 이용한 공압 인공근육 시스템의 인간-로봇 협력 제어장치 및 방법을 제공하고자 하는 것이다.

과제의 해결 수단

- [0032] 상기한 바와 같은 목적을 달성하기 위해, 본 발명에 따르면, 강화학습(reinforcement learning)을 이용한 공압 인공근육(Pneumatic Artificial Muscle ; PAM) 모듈의 인간-로봇 협력(human-robot cooperative) 제어장치에 있어서, 한 쌍의 공압 인공근육(PAM)이 케이블을 통해 폴리(pully)에 연결되는 것에 의해 서로 대립구성을 이루도록 형성되는 인공근육모듈; 상기 인공근육 모듈의 움직임을 제어하기 위한 제어스틱(control stick); 및 인공지능(AI)을 이용하여, 강화학습을 이용한 학습을 통해 상기 제어스틱에 가해지는 힘과 상기 제어스틱의 움직임에 대응하여 상기 인공근육모듈의 움직임을 제어하는 동시에, 상기 인공근육모듈의 움직임에 대응하는 힘을 상기 제어스틱에 전달하는 것에 의해 인간-로봇 상호작용(human-robot interaction)을 구현하는 처리가 수행되도록 이루어지는 제어부를 포함하여 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치가 제공된다.

- [0034] 여기서, 상기 제어장치는, 상기 제어장치의 본체를 형성하는 메인프레임; 상기 인공근육모듈의 각각의 공압 인공근육(PAM)을 구동하기 위해 상기 메인프레임에 설치되는 한 쌍의 PAM 액츄에이터(actuator); 상기 인공근육모

들의 각각의 공압 인공근육(PAM)의 움직임을 검출하기 위해 상기 메인프레임에 설치되는 한 쌍의 로드셀(loadcell); 상기 메인프레임에 회전 가능하게 설치되고 일단에 상기 풀리가 연결되는 샤프트; 상기 제어스틱의 움직임을 측정하기 위해 상기 제어스틱의 핸들과 상기 샤프트 사이에 설치되는 인터랙티브 로드셀(interactive loadcell); 상기 제어스틱의 움직임을 측정하기 위해 상기 샤프트의 타단에 설치되는 광학 로터리 엔코더(optical rotary encoder); 및 내부에 구비된 압력센서의 측정값에 근거하여 상기 제어스틱의 움직임에 따른 압력으로 압축공기를 상기 인공근육모듈의 각각의 공압 인공근육(PAM)에 전달하기 위해 상기 메인프레임에 설치되는 한 쌍의 비례 압력조절 밸브(proportional pressure regulator valve)를 더 포함하여 구성되는 것을 특징으로 한다.

[0036] 또한, 상기 제어부는, 이하의 수학식을 이용하여, 상기 제어스틱과 상기 인공근육모듈 사이의 인간-로봇 상호작용을 임피던스 모델로 나타내고,

$$\theta_m = \frac{K_h \tau_h}{M_d s^2 + C_d s + K_d}$$

[0038]

[0040] (여기서, θ_m 은 임피던스 모델의 출력위치(output position), M_d 는 관성(inertia) 파라미터 행렬, C_d 는 댐핑(damping) 파라미터 행렬, K_d 는 강성(stiffness) 파라미터 행렬, s 는 일반 비선형 방해(general nonlinear disturbance), K_h 는 인간 입력힘 이득(human input force gain), τ_h 는 인간 입력힘 이득 K_h 를 가지는 인간 제어노력(human control effort)을 각각 의미함.)

[0042] 상기 인공근육 모듈이 상기 임피던스 모델에 따라 동작하도록 토크 제어를 수행하는 내부루프(inner loop) 처리와, 상기 임피던스 모델에 대한 최적의 파라미터 값을 산출하는 외부루프(outer loop) 처리가 수행되도록 구성되는 것을 특징으로 한다.

[0044] 아울러, 상기 제어부는, 상기 외부루프(outer loop) 처리를 통해 산출된 상기 임피던스 모델에 대한 파라미터 값을 보상(reward)으로 적용하여 강화학습을 통해 최적의 파라미터 값을 업데이트 하는 처리가 수행되도록 구성되는 것을 특징으로 한다.

[0046] 더욱이, 상기 내부루프 처리는, 이하의 수학식을 이용하여, 상기 인공근육 모듈이 상기 임피던스 모델에 따라 동작하도록 하기 위한 제어토크(control torque)(u_c)를 산출하는 처리가 수행되도록 구성되는 것을 특징으로 한다.

$$u_c = \hat{\psi}(q) + K_s \sigma_h - u_r(t)$$

[0048]

[0050] (여기서, $u_r(t)$ 는 테일러 급수(Taylor series)의 고차항(high-order term)에 대한 견고성(robustness)을 제공하도록 특정된 강건화 신호(robustifying signal)이고, K_s 는 제어이득(control gain)이며, σ_h 는 슬라이딩모드 오

차(sliding mode error)이고, $\hat{\psi}(q)$ 는 $q = [\theta, \dot{\theta}, \ddot{\theta}_m, \ddot{\theta}_h, e_h, \dot{e}_h, \xi]^T$ 및 $\xi = \int_0^t e_h(\tau) d\tau$ 를 가지는 신경망(Neural Network ; NN)을 의미함)

[0052] 또한, 상기 외부루프 처리는, 관절각 오차(joint angle error) 및 인간-로봇 접촉력에 근거하여 상태공간(state space)을 형성하고 미리 정해진 간격으로 분할하는 처리가 수행되는 상태공간 및 이산화 처리단계; 상태에 따른 행동집합(actin set)을 정의하는 처리가 수행되는 행동집합 처리단계; 강화학습 알고리즘의 정책을 결정하는 처리가 수행되는 학습 알고리즘 처리단계; 상기 강화학습 알고리즘의 보상을 제공하고 상기 임피던스 모델의 파라미터를 업데이트 하기 위한 증분공식(incremental formular)을 도출하는 처리가 수행되는 업데이트 처리단계;

및 상기 증분공식에 근거하여 상기 임피던스 모델의 파라미터를 업데이트 하는 처리가 수행되는 파라미터 조정 단계를 포함하여 구성되는 것을 특징으로 한다.

[0054] 여기서, 상기 상태공간 및 이산화 처리단계는, 상기 광학 로터리 엔코더 및 상기 인터랙티브 로드셀을 통해 관절 각 오차(joint angle error) 및 인간-로봇 접촉력을 각각 측정하고 상태공간(state space)을 형성하며, 경험적 분할법(empirical segmentation method)을 이용하여 상기 관절 각 오차와 상기 인간-로봇 접촉력의 연속값(continuous value)을 M 및 N 간격(interval)으로 각각 나누어 상기 상태공간에 M×N 그리드를 형성하는 처리가 수행되도록 구성되는 것을 특징으로 한다.

[0056] 아울러, 상기 행동집합 처리단계는, 좌, 우 움직임(movement)에 따라 [1, 0], [0, 1]로 표현되는 두 그룹의 행동으로 상기 행동집합(actin set)을 정의하는 처리가 수행되도록 구성되는 것을 특징으로 한다.

[0058] 더욱이, 상기 학습 알고리즘 처리단계는, π 와 P를 각각 정책(policy) 및 상태전환 확률(state transition probability)이라 할 때, 이하의 수학적식을 이용하여 ρ -그리디(greedy) 정책을 강화학습 정책으로 적용하는 처리가 수행되도록 구성되는 것을 특징으로 한다.

$$\pi^{\rho} = \begin{cases} \arg \max_a Q(\theta, a) & \text{with probability } 1 - \rho \\ \frac{1}{A(\theta)} & \text{with probability } \rho \end{cases}$$

[0060]

[0062] (여기서, $Q(\theta, a)$ 는 상태 행동-값 함수(state action-value function) 이고, $A(\theta)$ 는 이산 행동집합(discrete action set)에 포함된 총 행동 수를 각각 의미함)

[0064] 또한, 상기 업데이트 처리단계는, 이하의 수학적식을 이용하여, 상태 행동-값 함수를 업데이트 하기 위한 증분공식(incremental formular)을 도출하는 처리가 수행되도록 구성되는 것을 특징으로 한다.

$$Q_{t+1}(\theta_t, a_t) = Q_t(\theta_t, a_t) + \alpha (R_{t+1} + \gamma Q_t(\theta_{t+1}, a_{t+1}) - Q_t(\theta_t, a_t))$$

[0066]

[0068] (여기서, t는 이산시간 단계(discrete-time step), α 는 단계크기(step size) 파라미터, γ 는 할인율(discount-rate) 파라미터, R_{t+1} 은 행동 a_t 를 취한 후 상태 θ_t 에서 상태 θ_{t+1} 로 전이(transition)를 가지는 시간 t+1의 보상을 각각 의미함)

[0070] 아울러, 상기 업데이트 처리단계에서, 상기 보상(R_{t+1})은 이하의 수학적식으로 나타내지는 것을 특징으로 한다.

$$R_{t+1} = r_{t+1} \mid a_t, \theta_t \rightarrow \theta_{t+1}$$

[0072]

[0074] (여기서, r_{t+1} 은 상태-행동쌍(state-action pair) (θ_t, a_t)의 평가(evaluation)에 해당하는 즉시보상(immediate reward)을 의미함)

[0076] 더욱이, 본 발명에 따르면, 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어방법에 있어서, 상기 에 기재된 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치를 이용하여 인공근육 모듈의 제어장치를 구현하는 단계; 및 상기 제어장치를 이용하여 상기 인공근육 모듈의 제어를 수행하는 단계를 포함하여 구성되는 것을 특징으로 하는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어방법이 제공된다.

발명의 효과

[0078] 상기한 바와 같이, 본 발명에 따르면, 케이블을 통해 폴리에 대립적 구성으로 연결된 한 쌍의 공압 인공근육(PAM)과, 인공근육의 구동을 위한 액추에이터 및 압축공기를 공급하는 비례압력조절 밸브와, 인공근육의 움직임을 제어하기 위한 제어스틱과, 각각의 공압 인공근육의 움직임을 측정하기 위한 로드셀과, 제어스틱의 움직임을 측정하기 위한 로드셀 및 광학 로터리 엔코더를 포함하여 구성되고, 미리 설정된 임피던스 모델에 따라 인간과

로봇의 상호작용을 측정하여 제어스틱의 움직임에 대응하는 힘을 인공근육에 전달하고 인공근육의 움직임에 대응하는 힘을 제어스틱에 전달하는 동시에, 강화학습을 이용한 학습을 통해 각 개인에게 적합한 최적의 모델을 결정하는 처리가 수행되도록 구성되는 강화학습을 이용한 공압 인공근육 시스템의 인간-로봇 협력 제어장치 및 방법이 제공됨으로써, 물리적 인간-로봇 상호작용(pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있고, 로봇이 규정된 동작모델에 따라 동작하도록 하여 상호작용을 단순화할 수 있는 동시에, 다양한 사용자에게 대하여도 용이하게 적용할 수 있다.

[0080] 또한, 본 발명에 따르면, 상기한 바와 같이 공압 인공근육(PAM) 모듈에 대하여 물리적 인간-로봇 상호작용(pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있도록 구성되는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법이 제공됨으로써, 인공근육의 움직임에 최대한 가깝게 인공근육의 움직임을 구현하는 데에만 중점을 두고 있는 것이 대부분이었던 종래기술의 공압 인공근육(PAM) 시스템 및 이를 이용한 로봇들의 한계를 해결할 수 있다.

도면의 간단한 설명

[0082] 도 1은 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치의 전체적인 구성을 개략적으로 나타내는 도면이다.

도 2는 도 1에 나타난 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치를 기반으로 한 인간-로봇 협동 시스템의 기본적인 개념을 설명하기 위해 인간-로봇 상호작용의 인간 손 모델을 개략적으로 나타내는 도면이다.

도 3은 도 2에 나타난 인간-로봇 상호작용 모델에 기반하여 인간-로봇 협업을 위해 이중루프를 가지는 제어체계의 전체적인 구조를 개략적으로 나타내는 도면이다.

도 4는 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어방법의 능동훈련 과정(active training phase)에서 임피던스 파라미터를 조정하는 처리과정의 전체적인 흐름을 개략적으로 나타내는 플로차트이다.

발명을 실시하기 위한 구체적인 내용

[0083] 이하, 첨부된 도면을 참조하여, 본 발명에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법의 구체적인 실시예에 대하여 설명한다.

[0085] 여기서, 이하에 설명하는 내용은 본 발명을 실시하기 위한 하나의 실시예일 뿐이며, 본 발명은 이하에 설명하는 실시예의 내용으로만 한정되는 것은 아니라는 사실에 유념해야 한다.

[0087] 또한, 이하의 본 발명의 실시예에 대한 설명에 있어서, 종래기술의 내용과 동일 또는 유사하거나 당업자의 수준에서 용이하게 이해하고 실시할 수 있다고 판단되는 부분에 대하여는, 설명을 간략히 하기 위해 그 상세한 설명을 생략하였음에 유념해야 한다.

[0089] 즉, 본 발명은, 후술하는 바와 같이, 인간근육의 움직임에 최대한 가깝게 인공근육의 움직임을 구현하는 데에만 중점을 두고 있는 것이 대부분이었던 종래기술의 공압 인공근육(PAM) 시스템 및 이를 이용한 로봇들의 한계를 해결하기 위해, 공압 인공근육(PAM) 모듈에 대하여 물리적 인간-로봇 상호작용(pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있도록 구성되는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법에 관한 것이다.

[0091] 아울러, 본 발명은, 후술하는 바와 같이, 물리적 인간-로봇 상호작용(pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있도록 하기 위해, 케이블을 통해 폴리에 대립적 구성으로 연결된 한 쌍의 공압 인공근육(PAM)과, 인공근육의 구동을 위한 액츄에이터 및 압축공기를 공급하는 비례압력조절 밸브와, 인공근육의 움직임을 제어하기 위한 제어스틱과, 각각의 공압 인공근육의 움직임을 측정하기 위한 로드셀과, 제어스틱의 움직임을 측정하기 위한 로드셀 및 광학 로터리 엔코더를 포함하여 구성되고, 미리 설정된 임피던스 모델에 따라 인간과 로봇의 상호작용을 측정하여 제어스틱의 움직임에 대응하는 힘을 인공근육에 전달하고 인공근육의 움직임에 대응하는 힘을 제어스틱에 전달하는 동시에, 강화학습을 이용한 학습을 통해 각 개인에게 적합한 최적의 모델을 결정하는 처리가 수행되도록 구성됨으로써, 로봇이 규정된 동작모델에 따라 동작하도록 하여 상호작용을 단순화할 수 있는 동시에, 다양한 사용자에게 적용할 수 있도록 구성되는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법에 관한 것이다.

- [0093] 계속해서, 도면을 참조하여, 본 발명에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법의 구체적인 내용에 대하여 설명한다.
- [0095] 먼저, 도 1을 참조하면, 도 1은 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치의 전체적인 구성을 개략적으로 나타내는 도면이다.
- [0097] 도 1에 나타난 바와 같이, 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치(10)는, 크게 나누어, 제어장치(10)의 본체를 형성하는 메인프레임(11)과, 한쌍의 공압 인공근육(PAM)이 케이블을 통해 풀리(pully)에 연결되는 것에 의해 서로 대립구성을 이루도록 형성되고 메인프레임(11)의 일측에 설치되어 공압으로 각각 수축/팽창 가능한 인공근육모듈(12)과, 각각의 공압 인공근육(PAM)을 구동하기 위해 메인프레임(11)의 일측에 함께 설치되는 한 쌍의 PAM 액츄에이터(actuator)(13)와, 각각의 공압 인공근육(PAM)의 움직임을 검출하기 위해 메인프레임(11)의 일측에 함께 설치되는 한 쌍의 로드셀(loadcell)(14)과, 인공근육 모듈(12)의 움직임을 제어하기 위해 샤프트를 통해 풀리와 연결되어 메인프레임(11)의 타측에 설치되는 제어스틱(control stick)(15)과, 제어스틱(15)의 움직임을 측정하기 위해 제어스틱(15)의 핸들과 샤프트 사이에 설치되는 인터랙티브 로드셀(interactive loadcell)(16)과, 제어스틱(15)의 움직임을 측정하기 위해 샤프트의 타단에 설치되는 광학 로터리 엔코더(optical rotary encoder)(17)와, 제어스틱(15)의 움직임에 따른 압력을 액츄에이터(13)를 통해 인공근육 모듈(12)에 전달하기 위한 한 쌍의 비례 압력조절 밸브(proportional pressure regulator valve)(18)를 포함하여 구성될 수 있다. 일 실시예로서, 상기 인공근육모듈(12)은, 한 쌍으로 구성되어 수축/팽창이 서로 반대되게 이루어질 수 있다. 즉, 한 개의 인공근육모듈(12)이 수축되면 다른 인공근육모듈(12)은 팽창될 수 있다.
- [0099] 여기서, 상기한 풀리는 알루미늄 프레임으로 이루어지고 베어링을 통해 중앙의 강철 샤프트에 회전 가능하도록 (swivel) 중앙의 강철 샤프트에 단단히 연결되어 있다.
- [0101] 또한, 도 1에 나타난 바와 같이, 스틸 샤프트의 타측에는 제어스틱(15)의 움직임을 측정하기 위해 광학 로터리 엔코더(17)가 설치되고, 각 PAM 액츄에이터(13)의 압축변위는 후술하는 바와 같이 하여 제어장치(10)의 기하학적 관계(geometrical relationship)에 의해 계산된다.
- [0103] 아울러, 상기한 제어장치(10)는, 압축공기를 각각 2개의 PAM에 공급하는 2개의 비례 압력조절 밸브(18)에 의해 제어스틱(15)을 통해 제어된 압력이 전달되도록 구성되며, 이때, 비례 압력조절 밸브(18) 내부에 구비된 압력센서에 의한 측정값에 근거하여 정확한 압력으로 제어되도록 구성될 수 있다.
- [0105] 더욱이, 상기한 제어장치(10)는, 도 1에 나타난 바와 같이, 스틸 샤프트와 제어스틱(15)의 핸들 사이에 인터랙티브 로드셀(16)이 장착되어 인간과 로봇의 상호작용을 측정할 수 있도록 구성될 수 있다.
- [0107] 또한, 상기한 제어장치(10)는, 도시되지는 않았으나, 후술하는 바와 같이 하여 인공지능(AI)을 이용한 강화학습을 통해 제어스틱(15)에 가해지는 힘과 움직임에 대응(비례)하여 인공근육모듈(12)의 움직임을 제어하는 동시에, 인공근육모듈(12)의 움직임에 대응(비례)하는 힘을 제어스틱(15)에 전달하는 것에 의해 인간-로봇 상호작용(human-robot interaction)을 구현하는 처리가 수행되도록 이루어지는 제어부를 더 포함하여 구성될 수 있다.
- [0109] 계속해서, 도 2 내지 도 4를 참조하여, 상기한 바와 같이 하여 구성되는 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치(10)의 구체적인 동작 및 작동원리에 대하여 설명한다.
- [0111] 먼저, 도 2를 참조하면, 도 2는 도 1에 나타난 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치(10)를 기반으로 한 인간-로봇 협동 시스템의 기본적인 개념을 설명하기 위해 인간-로봇 상호작용(human-robot interaction)의 인간 손 모델(human hand model)을 개략적으로 나타내는 도면이다.
- [0113] 아울러, 도 3을 참조하면, 도 3은 도 2에 나타난 인간-로봇 상호작용 모델에 기반하여 인간-로봇 협업을 위해 이중루프(double-loop)를 가지는 제어체계(control scheme)의 전체적인 구조를 개략적으로 나타내는 도면이다.
- [0115] 즉, 도 2에 있어서, 도 2는 간단한 1-자유도(degree of freedom ; DOF) 매스 스프링 댐퍼(mass-spring-damper) 모델을 기반으로 한 인간-로봇 협업(human-robot collaboration)의 기초적인 개념 설계를 나타내는 개념도이다.
- [0117] 여기서, 도 2에 있어서, m_h , k_h , c_h 는 인간 손-팔(human hand-arm)의 질량(mass), 감쇠(damping) 및 강성(stiffness)을 각각 나타내고, m_u , c_u 및 k_u 는 각각 제어장치(10)의 햅틱 장치의 질량, 감쇠 및 강성을 나타내는 것이다.

[0119] 더 상세하게는, 인간의 손은 그 쥐는 상태(grab condition)와 팔의 자세(posture)에 따라 강성과 감쇠가 변화하므로 비선형적인 강성과 감쇠를 가지고 있다.

[0121] 또한, 도 3에 나타난 제어구조에 있어서, e_d 에서 τ_h 까지의 인간 전달함수(human transfer function) $G(s)$ 는 이하의 [수학식 1]과 같이 가정될 수 있다.

[0123] [수학식 1]

$$G(s) = \frac{e_d}{\tau_h} = \frac{K_p + K_d s}{1 + k_e s}$$

[0124]

[0126] 여기서, 상기한 [수학식 1]에 있어서, K_p , K_d 및 k_e 는 각각 알려지지 않은 이득(unknown gain)을 의미하며, 이러한 파라미터(parameter)는 개인마다 다르고 특정 작업에 따라 달라지게 된다.

[0128] 즉, 인간이 능동적인 밀기-당기기 의식(active push-pull consciousness)을 가질 때 이에 영향을 받은 손은 특정한 움직임 능력(movement ability)을 가지게 되며, 로봇과 인간 사이의 조정되지 않은 움직임(uncoordinated movements)은 접촉력(contact force)을 행사하므로, 환경(environment)은 임피던스 특성(impedance characteristics)을 가진다.

[0130] 아울러, 로봇 컨트롤러의 설계는 접촉력을 줄이고 접촉력을 입력으로 하여 모션조정(motion adjustment)으로 반응해야 하므로, 따라서 두 물리적 시스템(physical system) 사이의 상호작용 모델(interactive model)(예를 들면, 환자와 로봇)은 도 3에 나타난 바와 같이 정해진(prescribed) 임피던스 특성을 가지게 된다.

[0132] 아울러, 도 3에 나타난 임피던스 모델에 있어서, 작업별 제어(task-specific control)에는 강제 폐쇄루프(force closed loop)가 없으며, 상호작용의 컴플라이언스(compliance)는 힘과 위치의 동적관계(dynamic relationship), 즉, 임피던스 모델의 특성으로 구현된다.

[0134] 또한, 인간-로봇 협력의 개발은 휴먼-인-더-루프(human-in-the-loop) 연구(investigation)에서 착안되었으며, 더 상세하게는, 첫째로, 인간에 대한 요구사항(requirement)을 제거하고(obviate) 로봇지향 모델(robot-oriented model)을 학습하기 위해 계산된(computed) 토크 컨트롤러(torque controller)가 사용된다.

[0136] 둘째로, 보조입력(assistive input)은 인간이 최소한의 노력(minimum effort)과 최고의 성과(best performance)로 정해진 작업(prescribed task)을 충족할 수 있도록 인간의 노력을 확대하기(magnify) 위해 적용된다.

[0138] 이에, 본 발명의 실시예에서는, 상기한 바와 같은 목표를 달성하기 위해, 도 3에 나타난 바와 같은 이중루프를 가지는 제어구조를 제안하였다.

[0140] 도 3에 있어서, 내측(inner one)은 수행된 작업에 대하여 어떠한 지식(knowledge)도 요구되지 않는 로봇지향 내부루프(robot oriented inner loop)이며, 외측은 인간과 로봇에 대한 역학(dynamics)과 작업정보(task information)를 포함하는 작업지향 외부루프(task oriented outer loop)이다.

[0142] 내측의 설계는 미확인된(unknown) 로봇이 인간에 의해 감지된 원하는 임피던스 모델처럼 행동하도록 하는 것이며, 따라서 인간은 임피던스 모델과 상호작용 하고, 내부루프에서 경로(trajecory)를 추적할 필요가 없으므로 외부루프에서 작업정보를 자유롭게 결합할 수 있다.

[0144] 로봇지향 내부루프 설계를 고려하면, 작업지향 외부루프로부터 학습된 임피던스 모델을 얻을 수 있으며, 이는, 추적오차(tracking error)를 최소화하고 사람이 최소한의 노력으로 작업을 수행할 수 있도록 도움을 줄 수 있다.

[0146] 아울러, 미확인된(unknown) 인간역학(human dynamics)이 고려되어야 하고, 인간-로봇 협력 시스템의 기대성능(expected performance)은 작업에 따라 달라지게 된다. 정해진(prescribed) 로봇 임피던스 모델은 이하의 [수학식 2]와 같이 나타낼 수 있다.

[0148] [수학식 2]

$$\theta_m = \frac{K_h \tau_h}{M_d s^2 + C_d s + K_d}$$

[0149]

[0151] 여기서, 상기한 [수학식 2]에 있어서, θ_m 은 원하는 임피던스 모델의 출력 위치(output position)를 나타내고, M_d , C_d 및 K_d 는 각각 원하는 관성(inertia), 감쇠 및 강성 파라미터 행렬을 나타내며, s 는 일반 비선형 방해(general nonlinear disturbance)를 나타내고, K_h 는 인간 입력힘 이득(human input force gain)을 나타내며, τ_h 는 인간 입력힘 이득 K_h 를 가지는 인간 제어노력(human control effort)을 각각 나타낸다.

[0153] 다음으로, 상기한 바와 같이 하여 구성되는 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치(10)의 제어설계(control design)에 대하여 설명한다.

[0155] 로봇의 고효율 추적을 보장하기 위해, 측정된 제어스틱(15)의 모션신호(motion signal)는 규정된(prescribed) 임피던스 모델 출력으로부터 원하는 궤적과 비교되고, 그 후, 이하의 위치제어에 의해 로봇에 대한 신호제어(signal control)가 생성된다.

[0157] 이러한 신호제어는 비례 압력조절 밸브를 보장하고(guarantee), 그 후 PAM(12)이 압축(compress)/확장(extend)함에 따라 폴리를 통해 와이어를 각각 당기기(pull)/밀기(push) 하여 컨트롤 스틱(15)도 움직이게 된다.

[0159] 내부루프 제어에 있어서, 시스템의 출력위치 θ 와 원하는 임피던스 모델의 출력위치 θ_m 사이의 힘 추적오차(force tracking error) e_h 는 이하의 [수학식 3]과 같이 정의될 수 있다.

[0161] [수학식 3]

$$e_h = \theta_m - \theta$$

[0162]

[0164] 여기서, 상기한 [수학식 3]에 있어서, 이는 궤도추종 오차(trjectory-following error)가 아니며, 따라서 대부분의 로봇 토크제어 작업과 달리, 이는 궤적추종 설계(trjectory-following design)가 아니라 모델추종 설계(model-following design)에 해당한다. 여기서는 작업정보가 요구되지 않으며, 모든 작업별 세부정보는 다음 단계에서 고려되고, 입력 제어토크(input control torque) τ 는 이하의 [수학식 4]와 같이 나타낼 수 있다.

[0166] [수학식 4]

$$\tau = u_c - K_h \tau_h$$

[0167]

[0169] 여기서, 상기한 [수학식 4]에 있어서, u_c 는 로봇이 상기한 [수학식 2]에 나타낸 임피던스 모델과 같이 동작하도록 하기 위한 제어토크(control torque)를 의미한다.

[0171] 다음으로, 로봇이 [수학식 2]에 나타낸 임피던스 모델과 같이 동작하도록 하기 위한 제어토크(control torque) u_c 를 설계해야 하며, 제어토크는 이하의 [수학식 5]와 같이 정의될 수 있다.

[0173] [수학식 5]

$$u_c = \hat{\psi}(q) + K_s \sigma_h - u_r(t)$$

$$\begin{cases} \hat{\psi}(q) = \hat{W}_1^T \phi(\hat{W}_2^T q) \\ \sigma_h = \dot{e}_h + \mu_1 e_h + \mu_2 \int_0^t e_h(\tau) d\tau \end{cases}$$

[0174]

[0176] 여기서, 상기한 [수학식 5]에 있어서, $u_r(t)$ 는 테일러 급수(Taylor series)의 고차항(high-order term)에 대한 견고성(robustness)을 제공하도록 특정된 강건화 신호(robustifying signal)이고, K_s 는 제어이득(control gain)이며, σ_h 는 슬라이딩모드 오차(sliding mode error)이고,

$$\hat{\psi}(q) \text{ 는 } q = [\theta, \dot{\theta}, \ddot{\theta}_m, \ddot{\theta}_m, e_h, \dot{e}_h, \xi]^T \text{ 및 } \xi = \int_0^t e_h(\tau) d\tau \text{ 를 가지는 신경망(Neural Network ; NN)이다.}$$

[0178] 아울러, 상기한 [수학식 5]에 있어서, NN 엘리먼트(element)는 미확인 로봇함수 $\psi(q)$ 설계를 이하의 [수학식 6]과 같이 보상(compensate)하는 데 사용된다.

[0180] [수학식 6]

$$\psi(q) = M(\ddot{\theta}_m + \mu_1 \dot{e}_h + \mu_1 e_h) + C(\dot{\theta}_m + \mu_1 e_h + \mu_2 \xi) + K\theta_m + \tau_d$$

[0181]

[0183] 더 상세하게는, 콤팩트 집합(compact set)에 대한 연속함수(continuous function) $\psi(q)$ 의 근사(approximation)인 2계층(two-layer) NN은 이하의 [수학식 7]과 같은 형태를 가진다.

[0185] [수학식 7]

$$\psi(q) = W_1^T \phi(W_2^T q) + \xi(q)$$

[0186]

[0188] 여기서, 상기한 [수학식 7]에 있어서, W_1 은 첫번째 계층 가중치 행렬(first-layer weight matrix)이고, W_2 는 두번째 계층 가중치 행렬(second-layer weight matrix)이며, ξ 은 NN 함수 근사오차(NN functional approximation error)를 각각 나타낸다.

[0190] 이때, 이상적인 가중치 벡터 W_1 및 W_2 는 알 수 없으며(unknown), 온라인으로 근사되고, 따라서 $\psi(q)$ 는 W_1 과 W_2 의 추정치(estimations)인 $\hat{W}_1$ 과 $\hat{W}_2$ 로 각각 근사된다.

[0192] 여기서, Z 를 이하의 [수학식 8]에 나타낸 바와 같이 정의하면, $\hat{Z}$ 는 W_1 과 W_2 대신에 $\hat{W}_1$ 과 $\hat{W}_2$ 와 동등하게 사용된다.

[0194] [수학식 8]

$$Z = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}$$

[0195]

[0197] 또한, 이상적인 NN 가중치에 $\|Z\| \leq \bar{Z}$ 의 경계가 있다고 가정하면, [수학식 5]에 의해 주어진 τ_c 를 가지는 [수학식 4]에 제안된 제어입력은 모델 추종오차(model following error) e_h 및 NN 가중치의 경계를 보장한다.

[0199] 아울러, NN 가중치에 대한 업데이트 규칙을 이하의 [수학식 9] 및 [수학식 10]에 나타낸 바와 같이 지정한다.

[0201] [수학식 9]

$$\dot{\hat{W}}_1 = F\hat{\phi}\sigma_h^T - F\hat{\phi}'\hat{W}_2^T q\sigma_h^T - kF\|\sigma_h\|\hat{W}_1$$

[0202]

[0204] [수학식 10]

$$\dot{\hat{W}}_2 = Gq(\hat{\phi}\hat{W}_1\sigma_h)^T - kG\|\sigma_h\|\hat{W}_2$$

[0205]

$$\hat{\phi} = \phi(\hat{W}_2^T q), \hat{\phi}' = d\phi(y)/dy|_{y=\hat{W}_2^T q}, F = F^T > 0, G = G^T > 0$$

[0207]

여기서, k 는 작은 양의 설계 파라미터(small positive design parameter)이며, 강건화항(robustifying term)은 이하의 [수학식 11]에 나타낸 바와 같다.

[0209] [수학식 11]

$$u_r(t) = -K_q(\|\hat{Z}\| + \bar{Z})$$

[0210]

[0212] 상기한 바와 같이, 내부루프의 위치 컨트롤러(position controller)는 임피던스모델 출력을 추적(track)하고, 접촉력(contact force)을 줄이기 위해, 접촉력을 입력으로 하고 모션조정(motion adjustment)으로 반응하며, 상호작용의 컴플라이언스(compliance)는 힘과 위치의 동적관계로 구현되고(embodied), 즉, 임피던스 모델의 특성이 온라인으로 업데이트 된다.

[0214] 또한, 외부루프 제어에 있어서, 상기한 임피던스 모델의 최적의 원하는 파라미터를 결정하기 위해 그러한 파라미터들의 개인화된 학습에 강화학습 알고리즘(reinforcement learning algorithm)이 적용된다. 학습 알고리즘의 목적은 사람의 수집된 상호작용 데이터를 이용하여 서로 다른 사람에 적합한 파라미터 조정을 위한 최적 전략을 찾기 위한 것이다.

[0216] 계속해서, 도 4를 참조하면, 도 4는 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어방법의 능동훈련과정(active training phase)에서 임피던스 파라미터를 조정하는 처리과정의 전체적인 흐름을 개략적으로 나타내는 플로차트이다.

[0218] 여기서, 강화학습이란, 기계학습(machine learning) 방법 중 하나로서 인공지능이 주어진 상태(state)에 대해 최적의 행동(action)을 선택하도록 하는 것으로 강화형 기계학습이라고도 불리며, 지도형/비지도형 기계학습에 이용되는 훈련데이터 대신 주어진 상태에 맞춘 행동의 결과에 대한 보상(reward)을 주는 것에 의해 인공지능의 성능을 향상시키는 것으로, 주로 게임이나 로봇제어 등에 적용되며, 이와 같이 학습을 위한 훈련데이터로 입력(주어진 상태)에 대한 출력(가장 적절한 행동)을 제공하는 대신에 행동의 좋고 나쁜 정도를 학습 알고리즘에게 알려주는 것을 보상(reward) 또는 강화(reinforcement)라 한다.

[0220] 아울러, 강화학습의 대상이 되는 컴퓨터 프로그램을 에이전트(agent) 라고도 하며, 에이전트는 주어진 상태(state)에서 자신이 취할 행동(action)을 표현하는 정책(policy)을 수립하고, 에이전트가 최대의 보상을 받을 수 있는 정책을 수립하도록 학습시키는 것이 강화학습의 목표이다.

[0222] 즉, 본 발명의 실시예에 있어서, 학습 알고리즘의 목적은, 수집된 그들 자신의 상호작용 데이터(interactive data)를 이용하여 다른 사람에게 적합한 파라미터 조정을 위한 최적의 전략(optimal strategy)을 찾는 것이며, 이를 위해, 로봇의 훈련을 위한 정의(definition)와 반복(iteration) 처리가 도 4에 나타낸 바와 같은 각 단계를 포함하여 구성될 수 있다.

[0224] 더 상세하게는, 도 4에 나타낸 처리흐름에 있어서, 첫째로, 상태-공간(state-space) 및 이산화(discretization) 단계에서, 관절각 오차(joint angle error) e_q (광학 로터리 엔코더로부터) 및 인간-로봇 접촉력 τ_h (인터랙티브 로드셀)가 측정되고, 시스템 상태(system state)로서 선택되어 상태공간(state space) $X = \{\theta, \tau_h\}$ 를 형성하기 위하여 선택된다.

[0226] 또한, 경험적 분할법(empirical segmentation method)을 이용하여 각도오차와 인간-로봇 접촉력의 연속값(continuous value)을 M 및 N 간격(interval)으로 각각 나누고, 이는 전체 상태공간을 나타내는데 사용되며, 배열(array)은 그리드(grid)의 넘버링(numbering)에 사용되고, 상태공간에는 $M \times N$ 그리드가 존재하게 된다.

[0228] 이때, 각각의 그리드는 $\{s_1, s_2\}$ 로 정의되며, 여기서 s_1 은 $[0, M]$ 범위의 각도오차(angle error) 간격번호

(interval number)를 나타내고, s_2 은 $[0, N]$ 범위의 인간-로봇 접촉력 간격번호를 나타낸다.

[0230] 둘째로, 상태 행동 집합(state action set) 단계에서, 행동집합(actin set) A는 대응하는 좌, 우 움직임(movement)에 따라 $[1, 0]$, $[0, 1]$ 로 표현되는 두 그룹의 행동으로 정의되며, 행동의 선택은 인간의 필요에 따라 목표상태(target state)가 주어졌을 때 상태전환(state transition)과 함께 최적의 경로(optimal route)에 대한 탐색이 된다.

[0232] 셋째로, 학습 알고리즘(lerning algorithm) 단계에서, π 와 P는 각각 정책(policy) 및 상태전환 확률(state transition probability)을 각각 나타내며, 강화학습은 의사결정에 사용되는 동일한 정책을 평가 또는 개선하고자 하는 정책에 따른 제어방법(on-policy control method)이다.

[0234] 여기서, 이러한 정책 방법으로 ρ -그리디(greedy) 정책이 적용될 수 있으며, 즉, 이는 $1-\rho$ 의 확률로 최대 추정 행동값(maximal estimated actin value)을 가지는 행동을 선택하나, 이하의 [수학식 12]에 나타낸 바와 같이, ρ 의 확률로 무작위로(random) 행동을 선택한다.

[0236] [수학식 12]

$$\pi^\rho = \begin{cases} \arg \max_a Q(\theta, a) & \text{with probability } 1 - \rho \\ \frac{1}{A(\theta)} & \text{with probability } \rho \end{cases}$$

[0237]

[0239] 여기서 $Q(\theta, a)$ 는 상태 행동-값 함수(state action-value function) 이고, $A(\theta)$ 는 이산 행동집합(discrete action set)에 포함된 총 행동 수이므로, 따라서 $1/A(\theta)$ 는 평균확률이 된다.

[0241] 넷째로, 상태 행동 업데이트 및 보상(reward) 단계에서, 상태 행동-값 함수 $Q(\theta, a)$ 및 보상(reward) R은, 이하의 [수학식 13]에 나타낸 바와 같이, 상태 행동-값 함수를 업데이트 하기 위한 증분공식(incremental formular)을 용이하게 고안(devise)할 수 있다.

[0243] [수학식 13]

$$Q_{t+1}(\theta_t, a_t) = Q_t(\theta_t, a_t) + \alpha (R_{t+1} + \gamma Q_t(\theta_{t+1}, a_{t+1}) - Q_t(\theta_t, a_t))$$

[0244]

[0246] 여기서 t는 이산시간 단계(discrete-time step), α 는 단계크기(step size) 파라미터, γ 는 할인율(discount-rate) 파라미터, R_{t+1} 은 이하의 [수학식 14]에 나타낸 바와 같이 표현되는 행동 a_t 를 취한 후 상태 θ_t 에서 상태 θ_{t+1} 로 전이(transition)를 가지는 시간 t+1의 보상을 각각 나타낸다.

[0248] [수학식 14]

$$R_{t+1} = r_{t+1} \mid a_t, \theta_t \rightarrow \theta_{t+1}$$

[0249]

[0251] 여기서, 상기한 [수학식 14]에 있어서, r_{t+1} 은 상태-행동쌍(state-action pair) (θ_t, a_t) 의 평가(evaluation)에 해당하는 즉시보상(immediate reward)이며, 이러한 정의는 작업 실행의 효과에 직접적인 영향을 미친다.

[0253] 더 상세하게는, 즉시보상 r_{t+1} 은 상수 -1로 설정되며, 이는, 상기한 바와 같이 초기 그리드(initial grid)로부터 목표 그리드(target grid)까지의 최단경로를 최적의 정책으로 선택하고 목표 그리드에 도달하기 전에 실패한 탐색(exploration)이기 때문이다. 또한, 즉시보상 값을 상수로 설정함으로써 전체적인 알고리즘의 처리과정을 단 순화할 수 있다.

[0255] 다섯째로, 임피던스 파라미터 업데이트 단계에서, 최종적으로 상기한 임피던스 모델 파라미터의 조정을 행한다.

[0257] 상기한 바와 같이, 본 발명의 실시예에 따르면, 인간-로봇 시스템의 상호작용 제어전략(interactive control strategy)은 수동적 협력(passive cooperation)과 능동적 협력(active cooperation)의 두 가지 훈련모드를 구

현하기 위한 상호작용 공간(interactive space)과 관절공간(joint space)의 제어구조를 포함하는 두 개의 루프에 의해 수행된다.

[0259] 여기서, 관절공간 제어전략에는 독립적 관절 분산제어 알고리즘(independent joint decentralized control algorithm)이 적용되고, 기계학습 중 인공지능이 주어진 상태(state)에 대하여 최적의 행동(action)을 선택하도록 하는 학습방법인 강화학습(reinforcement learning)에 기반한 개인화된 파라미터 학습방법(personalized parameter learning method)이 제안된다.

[0261] 또한, 상태변수(state variable)의 메시(mesh)는 인간-로봇 접촉력(human-robot contact force) 및 관절각 오차(joint angle error)를 이산화(discretizing) 하여 구성되고, 그것에 의해, 환자(patient)의 개인화된 특성이 학습 알고리즘에서 정량화된다(quantified).

[0263] 아울러, 강화학습을 통해 인간에게 적합한 개별화된 어드미턴스 파라미터(individualized admittance parameters)가 얻어지며, 이에, 본 발명에서 제시된 기술내용은 인간-로봇 협력(human-robot cooperation)의 개인적 적용(personal adaption) 및 능동 컴플라이언스(active compliance)에 기여할 수 있을 것으로 기대된다.

[0265] 따라서 상기한 바와 같이 하여, 본 발명의 실시예에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법을 구현할 수 있으며, 그것에 의해, 본 발명에 따르면, 케이블을 통해 폴리에 대립적 구성으로 연결된 한 쌍의 공압 인공근육(PAM)과, 인공근육의 구동을 위한 액추에이터 및 압축공기를 공급하는 비례압력조절 밸브와, 인공근육의 움직임을 제어하기 위한 제어스틱과, 각각의 공압 인공근육의 움직임을 측정하기 위한 로드셀과, 제어스틱의 움직임을 측정하기 위한 로드셀 및 광학 로터리 엔코더를 포함하여 구성되고, 미리 설정된 임피던스 모델에 따라 인간과 로봇의 상호작용을 측정하여 제어스틱의 움직임에 대응하는 힘을 인공근육에 전달하고 인공근육의 움직임에 대응하는 힘을 제어스틱에 전달하는 동시에, 강화학습을 이용한 학습을 통해 각 개인에게 적합한 최적의 모델을 결정하는 처리가 수행되도록 구성되는 강화학습을 이용한 공압 인공근육 시스템의 인간-로봇 협력 제어장치 및 방법이 제공됨으로써, 물리적 인간-로봇 상호작용(pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있고, 로봇이 규정된 동작모델에 따라 동작하도록 하여 상호작용을 단순화할 수 있는 동시에, 다양한 사용자에게 대하여도 용이하게 적용할 수 있다.

[0267] 아울러, 본 발명에 따르면, 상기한 바와 같이 공압 인공근육(PAM) 모듈에 대하여 물리적 인간-로봇 상호작용(pHRI)을 통하여 인공근육의 움직임을 보다 정확하고 효율적으로 제어할 수 있도록 구성되는 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법이 제공됨으로써, 인공근육의 움직임에 최대한 가깝게 인공근육의 움직임을 구현하는 데에만 중점을 두고 있는 것이 대부분이었던 종래기술의 공압 인공근육(PAM) 시스템 및 이를 이용한 로봇들의 한계를 해결할 수 있다.

[0269] 이상, 상기한 바와 같은 본 발명의 실시예를 통하여 본 발명에 따른 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 및 방법의 상세한 내용에 대하여 설명하였으나, 본 발명은 상기한 실시예에 기재된 내용으로만 한정되는 것은 아니며, 따라서 본 발명은, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 설계상의 필요 및 기타 다양한 요인에 따라 여러 가지 수정, 변경, 결합 및 대체 등이 가능한 것임은 당연한 일이라 하겠다.

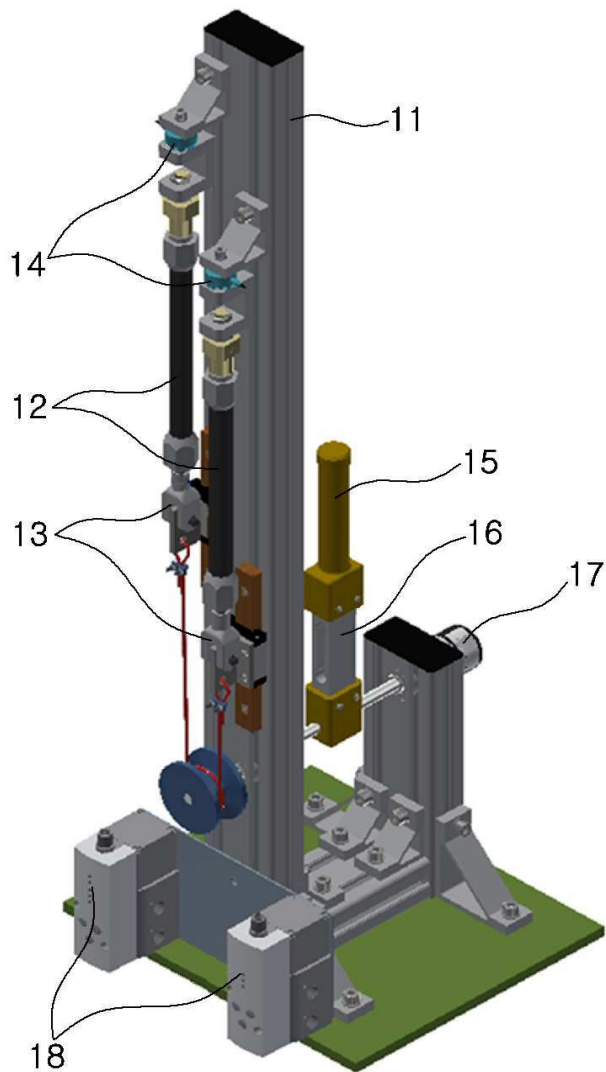
부호의 설명

- [0271]
- | | |
|---|----------------|
| 10. 강화학습을 이용한 공압 인공근육 모듈의 인간-로봇 협력 제어장치 | |
| 11. 메인프레임 | 12. 인공근육모듈 |
| 13. PAM 액추에이터 | 14. 로드셀 |
| 15. 제어스틱 | 16. 인터랙티브 로드셀 |
| 17. 광학 로터리 엔코더 | 18. 비례 압력조절 밸브 |

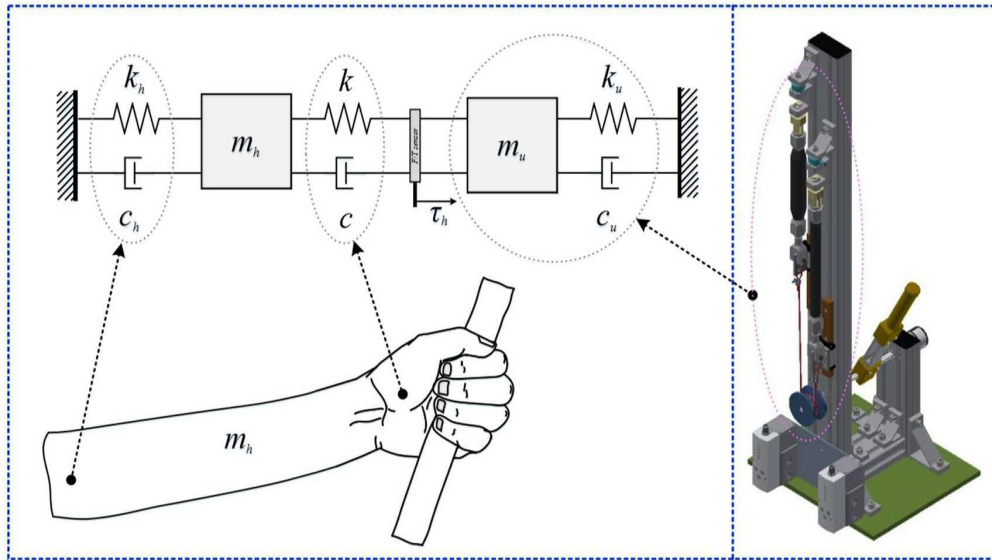
도면

도면1

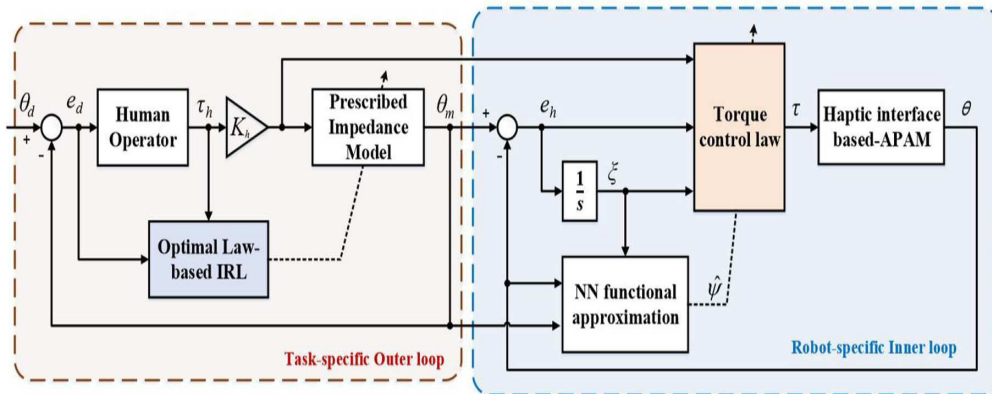
10



도면2



도면3



도면4

