

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일

2024 년 12 월 12 일 (12.12.2024)



WIPO | PCT



(10) 국제공개번호

WO 2024/253382 A1

(51) 국제특허분류:

G06V 40/20 (2022.01)

G06V 10/82 (2022.01)

G06V 10/778 (2022.01)

G06N 3/04 (2006.01)

04539 서울특별시 중구 을지로 65, SKT-타워, Seoul (KR). 장진성 (JANG, Jin Seong); 04539 서울특별시 중구 을지로 65, SKT-타워, Seoul (KR).

(21) 국제출원번호:

PCT/KR2024/007350

(22) 국제출원일:

2024 년 5 월 29 일 (29.05.2024)

(25) 출원언어:

한국어

(26) 공개언어:

한국어

(30) 우선권정보:

10-2023-0073992 2023 년 6 월 9 일 (09.06.2023) KR

(71) 출원인: 에스케이텔레콤 주식회사 (SK TELECOM CO., LTD.) [KR/KR]; 04539 서울특별시 중구 을지로 65 SK T-타워, Seoul (KR).

(72) 발명자: 마춘페이 (MA, Chun Fei); 04539 서울특별시 중구 을지로 65, SKT-타워, Seoul (KR). 윤영철 (YOON, Young Chul); 04539 서울특별시 중구 을지로 65, SKT-타워, Seoul (KR). 이병원 (LEE, Byeong Won);

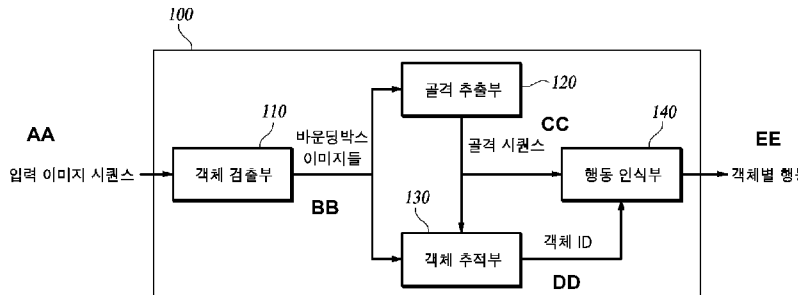
(74) 대리인: 이철희 (LEE, Chulhee); 06229 서울특별시 강남구 도곡로33길 26, 베리타스빌딩 2-4층, Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, CV, GH, GM,

(54) Title: APPARATUS AND METHOD FOR BEHAVIOR RECOGNITION BASED ON NOISE SKELETON SEQUENCE

(54) 발명의 명칭: 노이즈 골격 시퀀스 기반 행동 인식을 위한 장치 및 방법



110 ... Object detection unit
120 ... Skeleton extraction unit
130 ... Object tracking unit
140 ... Behavior recognition unit
AA ... Input image sequence
BB ... Bounding box images
CC ... Skeleton sequence
DD ... Object ID
EE ... Object-specific behavior

(57) Abstract: According to an aspect of the present invention, a computer-implemented method for training a behavior recognition model and an apparatus are provided, the method comprising the steps of: preparing a first skeleton sequence extracted from a first training image sequence by a skeleton extraction model, and behavior labels for the first skeleton sequence, wherein the first skeleton sequence includes first noise caused by the skeleton extraction model; and training a behavior recognition model on the basis of the first skeleton sequence and the behavior labels for the first skeleton sequence.

[다음 쪽 계속]



KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ,
UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ,
TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC,
ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

(57) 요약서: 본 발명의 일 측면에 의하면, 행동 인식 모델의 훈련을 위한 컴퓨터 구현 방법에 있어서, 골격 추출 모델이 제1 트레이닝 이미지 시퀀스로부터 추출한 제1 골격 시퀀스, 및 상기 제1 골격 시퀀스의 행동 레이블을 준비하는 단계 - 상기 제1 골격 시퀀스는 상기 골격 추출 모델에 의해 야기되는 제1 노이즈를 포함함 -; 및 상기 제1 골격 시퀀스 및 상기 제1 골격 시퀀스의 상기 행동 레이블을 기반으로 행동 인식 모델을 훈련하는 단계를 포함하는 방법 및 장치를 제공한다.

명세서

발명의 명칭: 노이즈 골격 시퀀스 기반 행동 인식을 위한 장치 및 방법

기술분야

- [1] 본 발명의 실시예들은 노이즈 골격 시퀀스 기반 행동 인식을 위한 장치 및 방법에 관한 것이다.

배경기술

- [2] 이하에 기술되는 내용은 단순히 본 실시예와 관련되는 배경 정보만을 제공할 뿐 종래기술을 구성하는 것이 아니다.
- [3] 최근 IT 기술의 발달에 따라, 카메라에 의해 촬영되는 영상으로부터 객체들의 행동을 인식하는 기술의 연구가 활발히 이루어지고 있다. 행동 인식 기술은 사람들의 행동 패턴 분석, 이상행동 감지, 침입자 감지 등 다양한 분야에서 이용된다.
- [4] 다양한 행동 인식 기술들 중 하나로서, 골격(skeleton)을 기반으로 객체의 행동을 인식하는 기술이 있다. 골격 기반 행동 인식 기술은 소정의 이미지 시퀀스 내 사람들의 골격을 기반으로 사람들의 자세 변화를 분석하고, 자세 변화로부터 사람들의 행동 유형들을 인식하는 기술이다.
- [5] 특히, 골격 기반 행동 인식 기술은 딥러닝 기반 행동 인식 모델을 이용하여 사람의 행동을 인식할 수 있다. 여기서, 행동 인식 모델은 사람의 골격 데이터로부터 사람의 행동 유형을 출력하도록 훈련되는 모델이다. 이때, 행동 인식 모델의 훈련 데이터에 대한 레이블로서, 작업자가 직접 애노테이션(annotation)한 골격 데이터가 이용된다. 골격 데이터의 애노테이션 결과는 그라운드 트루스(Ground Truth, GT)로 지칭될 수 있다.
- [6] 행동 인식 모델은 훈련 데이터와 유사한 분포를 갖는 입력 데이터에 대해 높은 인식 성능을 갖는다. 반면, 추론 단계에서 골격 데이터가 훈련 데이터와 다른 분포를 갖는 경우, 행동 인식 모델의 행동 인식 성능이 저하될 수 있다.
- [7] 구체적으로, 행동 인식 모델의 훈련에는 수동으로 애노테이션된 정확한 골격 데이터가 이용되지만, 행동 인식 모델의 추론에는 골격 추출 모델이 이미지로부터 추출한 부정확한 골격 데이터가 이용된다. 즉, 행동 인식 모델의 훈련 환경과 추론 환경이 다르다. 수동으로 애노테이션된 골격 데이터를 훈련한 행동 인식 모델은 골격 추출 모델에 의해 추출된 골격 데이터에 대한 행동 유형을 잘못 분류할 수 있다.
- [8] 특히, 골격 추출 모델은 관절들 중 움직임이 많은 관절들을 잘못 추출할 수 있다. 어느 행동 유형에 있어서, 움직임 변동이 큰 관절들은 검출되지 않거나 잘못된 위치로 검출될 수 있다. 예를 들면, 달리기 행동의 경우 다리와 팔의 움직임 변동이 크다. 골격 추출 모델이 달리기 행동에 관한 이미지들로부터 추출한 골격

데이터 중 다리 부위의 골격 데이터는 큰 오류를 포함할 수 있다. 따라서, 행동 인식 모델은 골격 추출 모델의 추출 결과에 대한 행동 유형을 잘못 분류할 수 있다.

[9] 이러한 문제점을 해결하기 위해서는, 행동 인식 모델이 다양한 자세에 따른 골격 데이터를 훈련해야 한다. 이는, 작업자가 더 많은 데이터를 애노테이션해야 하므로, 많은 비용과 시간이 소모된다.

[10] 따라서, 작업자의 애노테이션 외에 행동 인식 모델의 성능을 개선하기 위한 연구가 필요하다.

발명의 상세한 설명

기술적 과제

[11] 본 발명의 실시예들은, 적은 수의 그라운드 트루스로도 행동 인식 모델의 성능이 훈련 데이터와 다른 분포를 갖는 골격 데이터에 대해 저하되는 것을 방지하기 위한 훈련 방법 및 장치를 제공하는 데 주된 목적이 있다.

[12] 본 발명이 해결하고자 하는 과제들은 이상에서 언급한 과제들로 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제 해결 수단

[13] 본 발명의 일 측면에 의하면, 행동 인식 모델의 훈련을 위한 컴퓨터 구현 방법에 있어서, 골격 추출 모델이 제1 트레이닝 이미지 시퀀스로부터 추출한 제1 골격 시퀀스, 및 상기 제1 골격 시퀀스의 행동 레이블을 준비하는 단계 - 상기 제1 골격 시퀀스는 상기 골격 추출 모델에 의해 야기되는 제1 노이즈를 포함함 -; 및 상기 제1 골격 시퀀스 및 상기 제1 골격 시퀀스의 상기 행동 레이블을 기반으로 행동 인식 모델을 훈련하는 단계를 포함하는 방법을 제공한다.

[14] 본 실시예의 다른 측면에 의하면, 행동 인식 모델의 훈련을 위한 장치에 있어서, 명령어들을 저장하는 메모리; 및 적어도 하나의 프로세서를 포함하되, 상기 적어도 하나의 프로세서는 상기 명령어들을 실행함으로써, 골격 추출 모델이 제1 트레이닝 이미지 시퀀스로부터 추출한 제1 골격 시퀀스, 및 상기 제1 골격 시퀀스의 행동 레이블을 준비하고 - 상기 제1 골격 시퀀스는 상기 골격 추출 모델에 의해 야기되는 제1 노이즈를 포함함 -, 상기 제1 골격 시퀀스 및 상기 제1 골격 시퀀스의 상기 행동 레이블을 기반으로 행동 인식 모델을 훈련하는 장치를 제공한다.

[15] 본 실시예의 다른 측면에 의하면, 행동 인식을 위한 컴퓨터 구현 방법에 있어서, 입력 이미지 시퀀스로부터 추출된 입력 골격 시퀀스를 획득하는 단계; 및 행동 인식 모델을 이용하여 상기 입력 골격 시퀀스에 대한 행동 유형을 결정하는 단계를 포함하되, 상기 행동 인식 모델은 골격 추출 모델이 제1 트레이닝 이미지 시퀀스로부터 추출한 제1 골격 시퀀스, 및 상기 제1 골격 시퀀스의 행동 레이블을 준비하는 단계 - 상기 제1 골격 시퀀스는 상기 골격 추출 모델에 의해 야기되는 제1 노이즈를 포함함 -; 및 상기 제1 골격 시퀀스 및 상기 제1 골격 시퀀스의 상기 행

동 레이블을 기반으로 행동 인식 모델을 훈련하는 단계를 포함하여 훈련된 것인 방법을 제공한다.

발명의 효과

- [16] 이상에서 설명한 바와 같이 본 발명의 일 실시예에 의하면, 적은 수의 그라운드 트루스로도 행동 인식 모델의 성능이 훈련 데이터와 다른 분포를 갖는 골격 데이터에 대해 저하되는 것을 방지할 수 있다.
- [17] 본 개시의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [18] 도 1은 본 발명의 일 실시예에 따른 행동 인식 시스템의 구성도이다.
- [19] 도 2는 본 발명의 일 실시예에 따른 행동 인식 모델의 훈련을 설명하기 위한 도면이다.
- [20] 도 3은 본 발명의 다른 실시예에 따른 행동 인식 모델의 훈련을 설명하기 위한 도면이다.
- [21] 도 4는 본 발명의 다른 실시예에 따른 행동 인식 모델의 훈련을 설명하기 위한 도면이다.
- [22] 도 5는 본 발명의 일 실시예에 따른 제1 골격 데이터 및 제2 골격 데이터를 설명하기 위해 예시적으로 나타낸 도면이다.
- [23] 도 6은 본 발명의 일 실시예에 따른 행동 인식 모델의 훈련 방법의 순서도이다.
- [24] 도 7은 본 발명의 일 실시예에 따른 행동 인식 방법의 순서도이다.

발명의 실시를 위한 형태

- [25] 이하, 본 개시의 일부 실시예들을 예시적인 도면을 이용해 상세하게 설명한다. 각 도면의 구성 요소들에 참조 부호를 부가함에 있어서, 동일한 구성 요소들에 대해서는 비록 다른 도면 상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 개시를 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 개시의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.
- [26] 본 개시에 따른 실시예의 구성요소를 설명하는 데 있어서, 제1, 제2, i), ii), a), b) 등의 부호를 사용할 수 있다. 이러한 부호는 그 구성요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 부호에 의해 해당 구성요소의 본질 또는 차례나 순서 등이 한정되지 않는다. 명세서에서 어떤 부분이 어떤 구성요소를 '포함' 또는 '구비'한다고 할 때, 이는 명시적으로 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [27] 본 발명에 따른 장치 또는 방법의 각 구성요소는 하드웨어 또는 소프트웨어로 구현되거나, 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, 각 구성

요소의 기능이 소프트웨어로 구현되고 마이크로프로세서가 각 구성요소에 대응하는 소프트웨어의 기능을 실행하도록 구현될 수도 있다.

- [28] 도 1은 본 발명의 일 실시예에 따른 행동 인식 시스템의 구성도이다.
- [29] 도 1을 참조하면, 행동 인식 시스템(100)은 이미지들 내 객체들의 행동을 인식하는 시스템이다. 일 예로서, 행동 인식 시스템(100)은 입력 이미지 시퀀스 내 사람들을 식별 및 추적하고, 사람들의 앉는 행동, 서 있는 행동, 뛰는 행동, 쓰러지는 행동 등을 인식할 수 있다. 다른 예로서, 행동 인식 시스템(100)은 사람을 제외한 동물들에 대한 행동을 인식할 수도 있다.
- [30] 이를 위해, 행동 인식 시스템(100)은 객체 검출부(110), 골격 추출부(120), 객체 추적부(130) 및 행동 인식부(140)를 포함한다. 행동 인식 시스템(100)은 적어도 하나의 프로세서와 적어도 하나의 명령어를 포함하는 메모리를 포함하며, 적어도 하나의 프로세서에 의한 명령어 실행을 통해 객체 검출부(110), 골격 추출부(120), 객체 추적부(130) 및 행동 인식부(140)의 기능을 수행할 수 있다. 다른 실시예에서, 객체 검출부(110), 골격 추출부(120), 객체 추적부(130) 및 행동 인식부(140)가 개별 장치들에 의해 구현될 수 있다.
- [31] 객체 검출부(110)는 이미지 내 객체들을 검출한다. 여기서, 객체는 사람에 대한 그래픽 객체를 의미할 수 있다.
- [32] 구체적으로, 객체 검출부(110)는 객체 검출 모델을 이용하여 입력 이미지 시퀀스의 각 이미지 내 객체들을 포함하는 바운딩 박스들을 검출한다. 여기서, 하나의 바운딩 박스는 적어도 하나의 객체를 포함할 수 있다.
- [33] 한편, 객체 검출 모델은 객체들의 바운딩 박스를 생성하도록 미리 훈련된 딥러닝 모델로서, 합성곱 신경망 구조를 가질 수 있다. 일 예로서, 이미지들 내에 여러 사람들이 있는 경우, 객체 검출부(110)는 사람들 각각에 대한 바운딩 박스를 검출할 수 있다. 이후, 객체 검출부(110)는 이미지들로부터 바운딩 박스의 영역에 대응되는 바운딩박스 이미지들을 추출할 수 있다.
- [34] 골격 추출부(120)는 골격 추출 모델을 이용하여 바운딩박스 이미지들 내 객체의 골격 데이터를 추출한다. 골격 추출부(120)는 각 바운딩박스 이미지로부터 골격 데이터를 추출하고, 순차적 골격 데이터를 나타내는 골격 시퀀스를 출력한다.
- [35] 여기서, 객체의 골격 데이터는 사람의 관절들 및 관절들 간 링크들을 포함한다. 예를 들면, 골격 데이터는 사람의 머리, 어깨, 팔꿈치, 손목, 골반, 무릎 및 발목의 위치 좌표 및 연결 관계를 포함할 수 있다. 나아가, 골격 데이터는 눈, 코, 입, 귀, 목, 손끝, 몸통, 발끝 등의 관절 부위들을 더 포함할 수 있다.
- [36] 골격 시퀀스는 바운딩박스 이미지들로부터 순차적으로 추출된 골격 데이터의 집합이다. 골격 시퀀스는 N개의 프레임들로 구성되고, 각 프레임은 K개의 관절점들 및 관절점들 간 연결관계들을 포함할 수 있다. 일 예로서, 골격 시퀀스는 10개의 바운딩박스 이미지들 각각으로부터 추출된 10개의 관절점들 및 관절점들 간 연결관계들을 포함할 수 있다.

- [37] 골격 추출 모델은 주어진 이미지로부터 객체의 골격 데이터를 추출하도록 미리 훈련된 딥러닝 모델로서, 컨볼루션 신경망 구조를 가질 수 있다. 일 예에서, 골격 추출 모델은 주어진 이미지들에 대한 골격 데이터를 순차적으로 추출할 수 있다. 다른 예에서, 골격 추출 모델은 골격 데이터를 병렬적으로 추출할 수도 있다.
- [38] 골격 추출 모델은 입력 이미지 시퀀스에 기초하여 적응 또는 정밀 튜닝(fine-tuning)될 수 있다.
- [39] 객체 추적부(130)는 바운딩박스 이미지들 및 골격 시퀀스를 기반으로 객체를 추적하고, 객체에 ID를 할당한다.
- [40] 객체 추적부(130)는 추적 알고리즘을 이용하여 동일한 객체에 관한 바운딩박스 이미지들을 식별할 수 있다. 여기서, 추적 알고리즘은 객체 추적 기술 분야에서 자명하므로, 자세한 설명은 생략한다.
- [41] 행동 인식부(140)는 행동 인식 모델을 이용하여 골격 시퀀스에 대한 행동 유형을 결정하고, 객체 ID 및 행동 유형을 기반으로 객체별 행동 유형을 결정한다.
- [42] 여기서, 행동 유형은 사람이 앉는 행동, 서 있는 행동, 뛰는 행동, 쓰러지는 행동 등을 포함할 수 있다.
- [43] 행동 인식부(140)는 행동 인식 모델에 골격 시퀀스를 입력하고, 행동 인식 모델로부터 출력되는 행동 유형들에 대한 확률 분포를 획득하고, 행동 유형들에 대한 확률 분포를 기반으로 골격 시퀀스에 대한 행동 유형을 결정한다. 예를 들면, 행동 인식부(140)는 가장 높은 확률을 갖는 행동 유형을 결정할 수 있다.
- [44] 행동 인식 모델은 딥러닝 모델로서, 그래프 합성곱 신경망(Graphic Convolutional Network, GCN), ST-GCN(Spatial Temporal GCN), AS-GCN(Actional Structural GCN), CNN-LSTM(Convolution Neural Network-Long Short Term Memory) 등 다양한 구조를 가질 수 있다.
- [45] 행동 인식 모델은 골격 시퀀스를 입력 받고, 여러 행동 유형들에 대한 확률 분포를 출력하도록 훈련된다. 행동 인식 모델은 골격 시퀀스에 대응되는 행동 레이블과 행동 인식 모델로부터 출력되는 확률 분포 간 차이가 줄어들도록 훈련될 수 있다. 여기서, 하나의 행동 유형은 하나의 클래스를 가리키고, 행동 유형들에 대한 확률 분포는 신뢰도 점수(confidence score)를 가리킬 수 있다.
- [46] 특히, 훈련 데이터와 추론 데이터 간 차이에 의한 행동 인식 성능 저하를 방지하기 위해, 행동 인식 모델은 노이즈를 포함하는 골격 시퀀스를 기반으로 훈련된다. 노이즈를 포함하는 골격 시퀀스는 행동 인식 모델의 추론 데이터를 나타내는 입력 골격 시퀀스와 유사하다.
- [47] 구체적으로, 노이즈는 골격 추출 모델에 의해 야기되는 제1 노이즈, 또는 행동 유형들에 대한 관절들의 움직임 통계를 기반으로 생성된 제2 노이즈 중 적어도 하나를 포함한다.
- [48] 행동 인식 모델은 제1 노이즈를 포함하는 제1 골격 시퀀스 또는 제2 노이즈를 포함하는 제2 골격 시퀀스 중 적어도 하나를 기반으로 훈련된다. 제1 골격 시퀀스의 행동 레이블 및 제2 골격 시퀀스의 행동 레이블은 작업자에 의해 생성될 수

있다. 나아가, 행동 인식 모델은 수동으로 애노테이션된 그라운드 트루스에 해당하는 골격 시퀀스를 더 훈련할 수도 있다.

[49] 한편, 행동 인식 모델은 여러 명의 골격 시퀀스들을 순차적으로 처리할 수도 있고, 병렬적으로 처리할 수도 있다. 행동 인식 모델은 행동 인식 시스템(100)에 의해 훈련될 수도 있으나, 행동 인식 시스템(100)과는 다른 훈련 장치에 의해 훈련될 수 있다.

[50] 이처럼, 행동 인식 모델은 추론 대상인 입력 골격 시퀀스와 유사한 분포를 갖는 골격 시퀀스를 훈련하므로, 훈련 데이터와 추론 데이터 간 차이에 의한 행동 인식 성능이 개선될 수 있다.

[51] 도 2는 본 발명의 일 실시예에 따른 행동 인식 모델의 훈련을 설명하기 위한 도면이다.

[52] 도 2를 참조하면, 행동 인식 모델(220)은 골격 추출 모델(210)에 의해 추출된 제1 골격 시퀀스를 기반으로 훈련된다.

[53] 행동 인식 모델(220)은 적어도 하나의 프로세서 및 적어도 하나의 메모리를 포함하는 훈련 장치에 의해 훈련될 수 있다. 적어도 하나의 프로세서는 적어도 하나의 메모리에 저장된 명령어들을 실행함으로써, 행동 인식 모델(220)의 훈련 방법을 수행할 수 있다.

[54] 먼저, 훈련 장치는 제1 이미지 시퀀스를 준비한다. 제1 이미지 시퀀스는 한 사람에 대한 일련의 바운딩 박스 이미지들을 포함한다.

[55] 훈련 장치는 골격 추출 모델(210)을 이용하여 제1 트레이닝 이미지 시퀀스로부터 제1 골격 시퀀스를 추출한다. 골격 추출 모델(210)은 제1 트레이닝 이미지 시퀀스 내 사람의 관절들의 위치들 및 연결 관계들을 추출한다. 골격 추출 모델(210)은 제1 트레이닝 이미지 시퀀스 내 이미지들을 순차적으로 또는 병렬적으로 처리할 수 있다.

[56] 이때, 골격 추출 모델(210)이 추출한 제1 골격 시퀀스는 골격 추출 모델(210)에 의해 야기되는 제1 노이즈를 포함할 수 있다. 제1 골격 시퀀스는 골격 추출 모델(210)에 의해 야기되는 오류를 포함하는 부정확한 골격 데이터이다. 다만, 제1 골격 시퀀스는 그라운드 트루스에 비해 행동 인식 모델(220)의 추론 단계에서 골격 추출 모델(210)이 입력 이미지 시퀀스로부터 추출한 입력 골격 시퀀스와 유사할 수 있다.

[57] 골격 추출 모델(210)은 적절한 크기의 제1 노이즈를 야기하도록, 소정의 개수 이상의 파라미터들로 구성될 수 있다. 골격 추출 모델(210)의 파라미터 수를 나타내는 용량(capacity)이 너무 적으면, 제1 노이즈가 커질 수 있으므로, 골격 추출 모델(210)을 적절한 크기를 갖는다.

[58] 나아가, 훈련 장치는 제1 골격 시퀀스의 행동 레이블을 준비한다. 제1 골격 시퀀스의 행동 레이블은 제1 골격 시퀀스가 나타내는 행동 유형을 나타낸다. 일 예로서, 제1 골격 시퀀스의 행동 레이블은 원-핫 인코딩 벡터일 수 있다.

- [59] 훈련 장치는 제1 골격 시퀀스를 행동 인식 모델(220)에 입력하고, 제1 골격 시퀀스에 대한 행동 유형들의 확률 분포를 획득한다. 일 예로서, 행동 인식 모델(220)은 제1 골격 시퀀스를 입력 받고, 착석, 점프, 걷기 등의 행동 유형들에 대한 확률 분포를 출력한다.
- [60] 훈련 장치는 행동 인식 모델(220)에 의해 출력되는 확률 분포를 제1 골격 시퀀스의 행동 레이블과 비교하고, 비교 결과에 기초하여 행동 인식 모델(220)의 파라미터들을 업데이트한다. 구체적으로, 훈련 장치는 행동 인식 모델(220)에 의해 출력되는 확률 분포 및 제1 골격 시퀀스의 행동 레이블에 따라 제1 손실(Loss₁)을 계산하고, 제1 손실이 줄어들도록 행동 인식 모델(220)을 업데이트할 수 있다.
- [61] 여기서, 제1 손실을 계산하기 위해, 크로스 엔트로피(cross-entropy), 다중 클래스 로그 손실(multi class log loss), 이진 크로스 엔트로피(Binary cross-entropy), 또는 카테고리컬 크로스 엔트로피(Categorical cross-entropy) 등 분류 손실함수가 이용될 수 있다. 즉, 훈련 장치는 행동 인식 모델(220)에 의해 출력되는 확률 분포 및 제1 골격 시퀀스의 행동 레이블에 분류 손실함수를 적용하여 제1 손실을 계산할 수 있다.
- [62] 이처럼, 행동 인식 모델(220)이 골격 추출 모델(210)에 의해 제1 트레이닝 이미지 시퀀스로부터 추출된 제1 골격 시퀀스를 훈련하고, 추론 단계에서 골격 추출 모델(210)에 의해 입력 이미지 시퀀스로부터 추출되는 입력 골격 시퀀스로부터 행동을 추론하는 경우, 훈련 데이터와 추론 데이터 간 차이가 줄어들기 때문에 행동 인식 모델(220)의 행동 인식 성능이 개선될 수 있다.
- [63] 도 3은 본 발명의 다른 실시예에 따른 행동 인식 모델의 훈련을 설명하기 위한 도면이다.
- [64] 도 3을 참조하면, 행동 인식 모델(310)은 행동 유형들에 대한 관절들의 움직임 통계를 기반으로 생성된 제2 노이즈에 따른 제2 골격 시퀀스를 기반으로 훈련된다.
- [65] 여기서, 제2 노이즈는 행동 유형들에 대한 관절들의 움직임 통계를 기반으로 생성되는 노이즈이다. 제2 노이즈는 그라운드 트루스에 해당하는 참조 골격 시퀀스를 행동 인식 모델(310)의 추론 데이터와 유사해지도록 하기 위한 것이다.
- [66] 구체적으로, 어느 행동 유형에 대해 관절별로 움직임 정도가 다르다. 예를 들면, 사람의 달리기 행동에서, 사람의 몸통에 비해 팔과 다리가 많이 움직인다. 또한, 사람의 점프 행동 및 착석 행동에서는 다리가 많이 움직인다.
- [67] 행동 인식 모델(310)의 추론에 이용되는 골격 추출 모델은 움직임이 많은 관절들에 대해서는 검출 성능이 악화될 수 있다. 예를 들면, 사람의 달리기 행동을 추론하기 위해, 골격 추출 모델에 의해 추출된 다리 관절 위치들은 실제 다리 관절 위치와는 다를 수 있으며, 다른 부위의 관절들에 비해 오차가 클 수 있다.
- [68] 행동 인식 모델(310)이 그라운드 트루스에 해당하는 참조 골격 시퀀스를 훈련한 경우, 사람의 부위별로 다른 오차를 포함하는 골격 시퀀스에 대해 행동 유형을 잘못 분류할 가능성이 높다.

- [69] 따라서, 훈련 장치는 행동 유형들에 대한 관절들의 움직임 통계에 따라 변형된 제2 골격 시퀀스를 기반으로 행동 인식 모델(310)을 훈련한다.
- [70] 먼저, 훈련 장치는 제2 이미지 시퀀스를 준비한다. 제2 이미지 시퀀스는 도 2의 제1 이미지 시퀀스와 같을 수도 있고, 다를 수도 있다.
- [71] 이후, 라벨링에 의해 제2 이미지 시퀀스에 대한 참조 골격 시퀀스가 생성된다. 참조 골격 시퀀스는 제2 이미지 시퀀스 내 사람의 관절점들 및 관절점들 간 링크들을 포함한다. 참조 골격 시퀀스는 작업자에 의해 수동으로 생성되는 그라운드 트루스이다. 그렇지 않으면, 참조 골격 시퀀스는 라벨링 모델에 의해 생성될 수 있다. 참조 골격 시퀀스는 사람의 관절점들 및 관절점들 간 링크들을 정확하게 나타낸 것이다.
- [72] 훈련 장치는 행동 유형들에 대한 관절들의 움직임 통계를 미리 저장한다. 행동 유형들에 대한 관절들의 움직임 통계는 각 행동 유형에 대해 각 관절의 움직임 변화량을 포함한다.
- [73] 훈련 장치는 행동 유형들에 대한 관절들의 움직임 통계를 기반으로 제2 노이즈를 생성한다. 일 예로서, 훈련 장치는 어느 관절의 움직임 정도에 따른 범위를 설정하고, 설정된 범위 내에서 임의로 해당 관절의 이동량 및 이동방향을 제2 노이즈로 생성할 수 있다.
- [74] 훈련 장치는 참조 골격 시퀀스에 제2 노이즈를 부가하여 제2 골격 시퀀스를 생성한다. 훈련 장치는 참조 골격 시퀀스의 각 관절 위치들을 제2 노이즈에 따라 이동시킴으로써 제2 골격 시퀀스를 생성할 수 있다.
- [75] 이후, 훈련 장치는 제2 골격 시퀀스를 행동 인식 모델(310)에 입력하고, 제2 골격 시퀀스에 대한 행동 유형들의 확률 분포를 획득한다. 훈련 장치는 행동 인식 모델(310)에 의해 출력되는 확률 분포를 제2 골격 시퀀스의 행동 레이블과 비교하고, 비교 결과에 기초하여 행동 인식 모델(310)의 파라미터들을 업데이트한다. 훈련 장치는 행동 인식 모델(310)에 의해 출력되는 확률 분포 및 제2 골격 시퀀스의 행동 레이블에 따라 제2 손실(Loss₂)을 계산하고, 제2 손실이 줄어들도록 행동 인식 모델(310)을 업데이트할 수 있다. 제2 손실을 계산하기 위한 손실함수는 전술한 분류 손실함수들 중 하나가 이용될 수 있다.
- [76] 이처럼, 행동 인식 모델(310)이 행동 유형들에 대한 관절들의 움직임 통계를 기반으로 생성된 제2 골격 시퀀스를 훈련하는 경우, 훈련 데이터와 추론 데이터 간 차이를 줄여 행동 인식 모델(310)의 행동 인식 성능을 개선할 수 있다.
- [77] 나아가, 제2 골격 시퀀스는 참조 골격 시퀀스로부터 증강되는 것이므로, 행동 인식 모델(310)의 훈련 데이터를 구축하는 비용이 절감될 수 있다.
- [78] 도 4는 본 발명의 다른 실시예에 따른 행동 인식 모델의 훈련을 설명하기 위한 도면이다.
- [79] 도 4를 참조하면, 행동 인식 모델(410)은 골격 추출 모델에 의해 추출되는 제1 골격 시퀀스, 및 행동 유형들에 대한 움직임 통계를 기반으로 생성되는 제2 노이즈를 포함하는 제2 골격 시퀀스를 모두 훈련할 수 있다.

- [80] 여기서, 제1 골격 시퀀스 및 제2 골격 시퀀스는 각각 도 2 및 도 3에서 설명된 바와 같다.
- [81] 훈련 장치는 제1 골격 시퀀스 및 제1 골격 시퀀스의 행동 레이블을 기반으로 제1 손실을 계산하고, 제2 골격 시퀀스 및 제2 골격 시퀀스의 행동 레이블을 기반으로 제2 손실을 계산하고, 제1 손실과 제2 손실의 가중합을 행동 인식 모델(410)의 전체 손실로 계산한다. 제1 손실과 제2 손실의 가중 비율을 나타내는 τ_1 및 τ_2 는 휴리스틱으로 설정될 수 있다.
- [82] 훈련 장치는 전체 손실이 감소하도록 행동 인식 모델(410)의 파라미터들을 업데이트한다.
- [83] 도 5는 본 발명의 일 실시예에 따른 제1 골격 데이터 및 제2 골격 데이터를 설명하기 위해 예시적으로 나타낸 도면이다.
- [84] 도 5를 참조하면, 사람이 걷는 동작을 캡처한 트레이닝 이미지(510)가 도시되어 있다.
- [85] 행동 인식 모델의 추론 단계에서, 골격 추출 모델(210)이 입력 이미지 시퀀스로부터 추출한 입력 골격 시퀀스는 노이즈를 포함한다. 특히, 움직임이 많은 관절에 대해 상대적으로 큰 오차를 포함할 가능성이 높다. 행동 인식 모델의 추론 성능을 개선하기 위해, 이러한 노이즈를 포함하는 골격 시퀀스를 훈련 데이터로 이용한다.
- [86] 먼저, 훈련 장치는 트레이닝 이미지(510)를 골격 추출 모델(210)에 입력하여 제1 골격 데이터(520)를 추출한다. 여기서, 제1 골격 데이터(520)는 골격 추출 모델(210)에 의해 야기되는 제1 노이즈를 포함한다. 제1 골격 데이터(520)는 제1 노이즈에 따라 하체 관절들이 어긋나 있다.
- [87] 훈련 장치는 트레이닝 이미지(510)를 수동으로 애노테이션한 참조 골격 데이터(530)를 획득하고, 참조 골격 데이터(530)에 제2 노이즈를 부가하여 제2 골격 데이터(540)를 생성한다. 통계적으로 걷기 동작에서 다리의 움직임이 많으므로, 참조 골격 데이터(530)의 상체 관절들에 비해 하체 관절들에 큰 노이즈가 부가된다. 이에 따라, 제2 골격 데이터(540)는 하체 관절들이 많이 벌어져 있다.
- [88] 행동 인식 모델의 추론 단계에서, 행동 인식 모델에 입력되는 골격 시퀀스는 골격 추출 모델(210)에 의해 추출되는 것이므로, 제1 골격 데이터(520) 및 제2 골격 데이터(540)와 같은 오차들을 포함할 확률이 높다.
- [89] 참조 골격 데이터(530)를 훈련한 행동 인식 모델은 제1 골격 데이터(520) 및 제2 골격 데이터(540)와 유사한 골격 데이터에 대해 낮은 인식 성능을 보이는 반면, 제1 골격 데이터(520) 및 제2 골격 데이터(540)를 훈련한 행동 인식 모델은 제1 골격 데이터(520) 및 제2 골격 데이터(540)와 유사한 골격 데이터에 대해 높은 인식 성능을 보인다.
- [90] 도 6은 본 발명의 일 실시예에 따른 행동 인식 모델의 훈련 방법의 순서도이다.

- [91] 도 6을 참조하면, 행동 인식 모델을 훈련하기 위한 훈련 장치는 골격 추출 모델이 제1 트레이닝 이미지 시퀀스로부터 추출한 제1 골격 시퀀스, 및 제1 골격 시퀀스의 행동 레이블을 준비한다(S610).
- [92] 여기서, 제1 골격 시퀀스는 골격 추출 모델에 의해 야기되는 제1 노이즈를 포함한다.
- [93] 또한, 훈련 장치는 제2 트레이닝 이미지 시퀀스의 참조 골격 시퀀스에 제2 노이즈를 부가하여 생성한 제2 골격 시퀀스 및 제2 골격 시퀀스의 행동 레이블을 준비한다(S620).
- [94] 여기서, 제2 노이즈는 행동 유형들에 대한 관절들의 움직임 통계를 기반으로 생성된 것이다.
- [95] 이후, 훈련 장치는 제1 골격 시퀀스, 제1 골격 시퀀스의 행동 레이블, 제2 골격 데이터 및 제2 골격 데이터의 행동 레이블에 더 기초하여 행동 인식 모델을 훈련한다(S630).
- [96] 구체적으로, 훈련 장치는 행동 인식 모델이 제1 골격 시퀀스로부터 결정한 행동 유형과 제1 골격 시퀀스의 행동 레이블을 비교하여 제1 손실을 계산한다. 훈련 장치는 행동 인식 모델이 제2 골격 시퀀스로부터 결정한 행동 유형과 제2 골격 시퀀스의 행동 레이블을 비교하여 제2 손실을 계산한다. 훈련 장치는 제1 손실과 제2 손실의 가중합을 전체 손실로 계산하고, 전체 손실이 줄어드는 방향으로 행동 인식 모델을 업데이트한다.
- [97] 도 7은 본 발명의 일 실시예에 따른 행동 인식 방법의 순서도이다.
- [98] 이하에서, 행동 인식 장치는 도 2에서 행동 인식부(140)를 가리킨다.
- [99] 도 7을 참조하면, 행동 인식 장치는 입력 이미지 시퀀스로부터 추출된 입력 골격 시퀀스를 획득한다(S710).
- [100] 여기서, 입력 골격 시퀀스는 골격 추출 모델에 의해 추출된 것이며, 골격 추출 모델은 행동 인식 모델의 훈련 데이터를 생성하는 데 이용된 것이다.
- [101] 행동 인식 장치는 행동 인식 모델을 이용하여 입력 골격 시퀀스에 대한 행동 유형을 결정한다(S720).
- [102] 구체적으로, 행동 인식 모델은 입력 골격 시퀀스가 행동 유형들에 속할 확률 분포를 출력하고, 행동 인식 장치는 입력 골격 시퀀스가 속할 확률이 가장 높은 행동 유형을 입력 골격 시퀀스의 행동 유형으로 결정한다.
- [103] 여기서, 행동 인식 모델은 도 2 내지 도 6에서 설명된 훈련 방법들 중 하나에 따라 훈련된 모델이다.
- [104] 본 명세서에 설명되는 시스템들 및 기법들의 다양한 구현예들은, 디지털 전자 회로, 집적회로, FPGA(field programmable gate array), ASIC(application specific integrated circuit), 컴퓨터 하드웨어, 펌웨어, 소프트웨어, 및/또는 이들의 조합으로 실현될 수 있다. 이러한 다양한 구현예들은 프로그래밍가능 시스템 상에서 실행 가능한 하나 이상의 컴퓨터 프로그램들로 구현되는 것을 포함할 수 있다. 프로그래밍가능 시스템은, 저장 시스템, 적어도 하나의 입력 디바이스, 그리고 적

어도 하나의 출력 디바이스로부터 데이터 및 명령들을 수신하고 이들에게 데이터 및 명령들을 전송하도록 결합되는 적어도 하나의 프로그래밍가능 프로세서(이것은 특수 목적 프로세서일 수 있거나 혹은 범용 프로세서일 수 있음)를 포함한다. 컴퓨터 프로그램들(이것은 또한 프로그램들, 소프트웨어, 소프트웨어 애플리케이션들 혹은 코드로서 알려져 있음)은 프로그래밍가능 프로세서에 대한 명령어들을 포함하며 "컴퓨터가 읽을 수 있는 기록매체"에 저장된다.

- [105] 컴퓨터가 읽을 수 있는 기록매체는, 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 이러한 컴퓨터가 읽을 수 있는 기록매체는 ROM, CD-ROM, 자기 테이프, 플로피디스크, 메모리 카드, 하드 디스크, 광자기 디스크, 스토리지 디바이스 등의 비휘발성(non-volatile) 또는 비일시적인(non-transitory) 매체일 수 있으며, 또한 데이터 전송 매체(data transmission medium)와 같은 일시적인(transitory) 매체를 더 포함할 수도 있다. 또한, 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다.
- [106] 본 명세서의 흐름도/타이밍도에서는 각 과정들을 순차적으로 실행하는 것으로 기재하고 있으나, 이는 본 개시의 일 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것이다. 다시 말해, 본 개시의 일 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 개시의 일 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 흐름도/타이밍도에 기재된 순서를 변경하여 실행하거나 각 과정들 중 하나 이상의 과정을 병렬적으로 실행하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이므로, 흐름도/타이밍도는 시계열적인 순서로 한정되는 것은 아니다.
- [107] 이상의 설명은 본 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 실시예들은 본 실시예의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.
- [108] CROSS-REFERENCE TO RELATED APPLICATION
- [109] 본 특허출원은, 본 명세서에 그 전체가 참고로서 포함되는, 2023년 06월 09일에 한국에 출원한 특허출원번호 제10-2023-0073992호에 대해 우선권을 주장한다.

청구범위

- [청구항 1] 적어도 하나의 프로세서를 포함하는 훈련 장치에 의해 수행되는, 행동 인식 모델의 훈련을 위한 방법에 있어서,
 골격 추출 모델이 제1 트레이닝 이미지 시퀀스로부터 추출한 제1 골격 시퀀스, 및 상기 제1 골격 시퀀스의 행동 레이블을 준비하는 단계 - 상기 제1 골격 시퀀스는 상기 골격 추출 모델에 의해 야기되는 제1 노이즈를 포함 -; 및
 상기 제1 골격 시퀀스 및 상기 제1 골격 시퀀스의 상기 행동 레이블을 기반으로 행동 인식 모델을 훈련하는 단계를 포함하는 방법.
- [청구항 2] 제1항에 있어서,
 제2 트레이닝 이미지 시퀀스의 참조 골격 시퀀스에 제2 노이즈를 부가하여 생성한 제2 골격 시퀀스 및 상기 제2 골격 시퀀스의 행동 레이블을 준비하는 단계 - 상기 제2 노이즈는 행동 유형들에 대한 관절들의 움직임 통계를 기반으로 생성된 것임 -;
 를 더 포함하는 방법.
- [청구항 3] 제2항에 있어서,
 상기 행동 인식 모델을 훈련하는 단계는,
 상기 제2 골격 시퀀스 및 상기 제2 골격 시퀀스의 행동 레이블에 더 기초하여 상기 행동 인식 모델을 훈련하는 단계를 포함하는 방법.
- [청구항 4] 행동 인식 모델의 훈련을 위한 장치에 있어서,
 명령어들을 저장하는 메모리; 및
 적어도 하나의 프로세서를 포함하되,
 상기 적어도 하나의 프로세서는 상기 명령어들을 실행함으로써,
 골격 추출 모델이 제1 트레이닝 이미지 시퀀스로부터 추출한 제1 골격 시퀀스, 및 상기 제1 골격 시퀀스의 행동 레이블을 준비하고 - 상기 제1 골격 시퀀스는 상기 골격 추출 모델에 의해 야기되는 제1 노이즈를 포함함 -,
 상기 제1 골격 시퀀스 및 상기 제1 골격 시퀀스의 상기 행동 레이블을 기반으로 행동 인식 모델을 훈련하는 장치.
- [청구항 5] 행동 인식을 위한 컴퓨터 구현 방법에 있어서,
 입력 이미지 시퀀스로부터 추출된 입력 골격 시퀀스를 획득하는 단계; 및
 행동 인식 모델을 이용하여 상기 입력 골격 시퀀스에 대한 행동 유형을 결정하는 단계
 를 포함하되,
 상기 행동 인식 모델은,

골격 추출 모델이 제1 트레이닝 이미지 시퀀스로부터 추출한 제1 골격 시퀀스, 및 상기 제1 골격 시퀀스의 행동 레이블을 준비하는 단계 - 상기 제1 골격 시퀀스는 상기 골격 추출 모델에 의해 야기되는 제1 노이즈를 포함함 -; 및

상기 제1 골격 시퀀스 및 상기 제1 골격 시퀀스의 상기 행동 레이블을 기반으로 행동 인식 모델을 훈련하는 단계를 포함하여 훈련된 것인 방법.

[청구항 6]

제5항에 있어서,

상기 행동 인식 모델은,

제2 트레이닝 이미지 시퀀스의 참조 골격 시퀀스에 제2 노이즈를 추가하여 생성한 제2 골격 시퀀스 및 상기 제2 골격 시퀀스의 행동 레이블을 준비하는 단계 - 상기 제2 노이즈는 행동 유형들에 대한 관절들의 움직임 통계를 기반으로 생성된 것임 -; 및

상기 제2 골격 시퀀스 및 상기 제2 골격 시퀀스의 행동 레이블에 더 기초하여 상기 행동 인식 모델을 훈련하는 단계를 더 포함하여 훈련된 것인 방법.

[청구항 7]

제6항에 있어서,

상기 입력 골격 시퀀스는,

상기 골격 추출 모델에 의해 추출된 것인 방법.

[청구항 8]

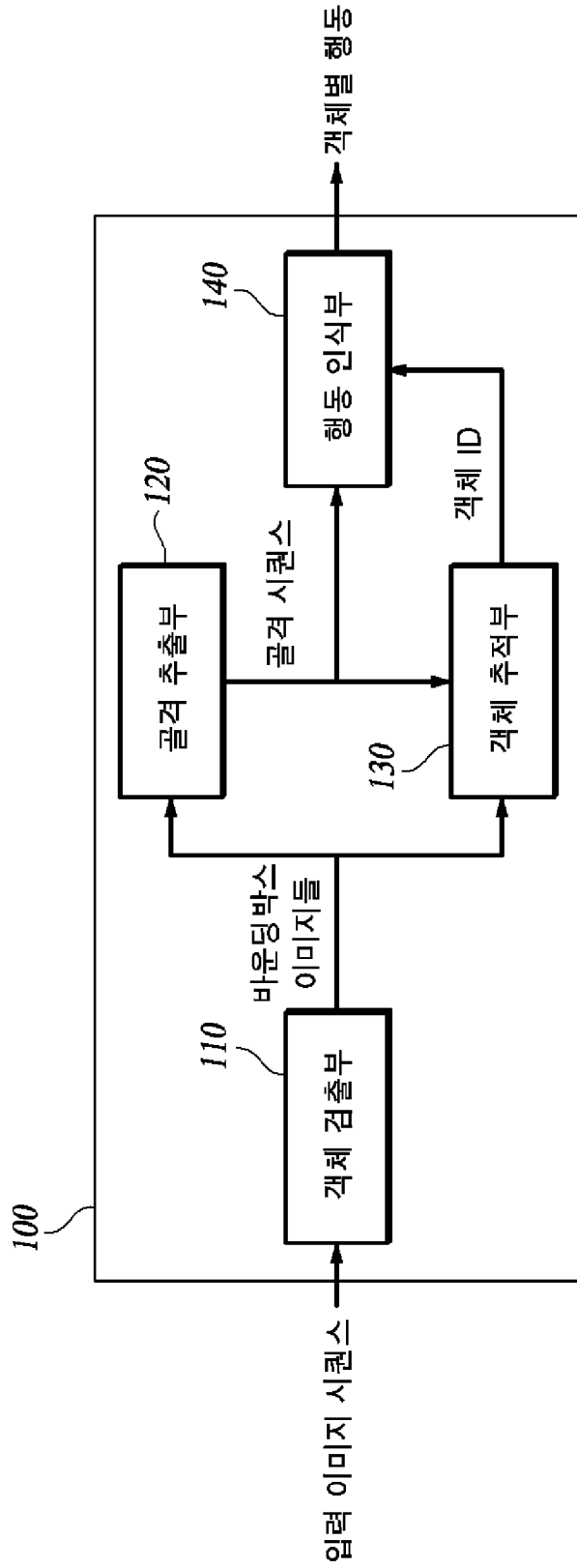
명령어가 저장된, 컴퓨터로 읽을 수 있는 기록매체로서, 상기 명령어는 상기 컴퓨터에 의해 실행될 때 상기 컴퓨터로 하여금,

골격 추출 모델이 제1 트레이닝 이미지 시퀀스로부터 추출한 제1 골격 시퀀스, 및 상기 제1 골격 시퀀스의 행동 레이블을 준비하는 단계 - 상기 제1 골격 시퀀스는 상기 골격 추출 모델에 의해 야기되는 제1 노이즈를 포함함 -; 및

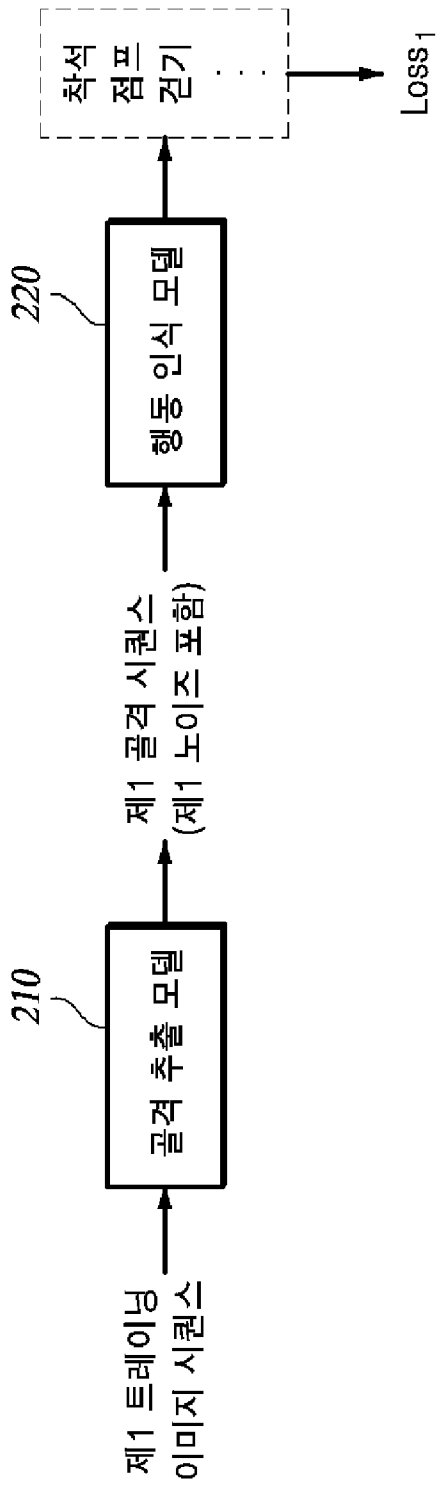
상기 제1 골격 시퀀스 및 상기 제1 골격 시퀀스의 상기 행동 레이블을 기반으로 행동 인식 모델을 훈련하는 단계를

를 실행하도록 하는, 컴퓨터로 읽을 수 있는 기록매체.

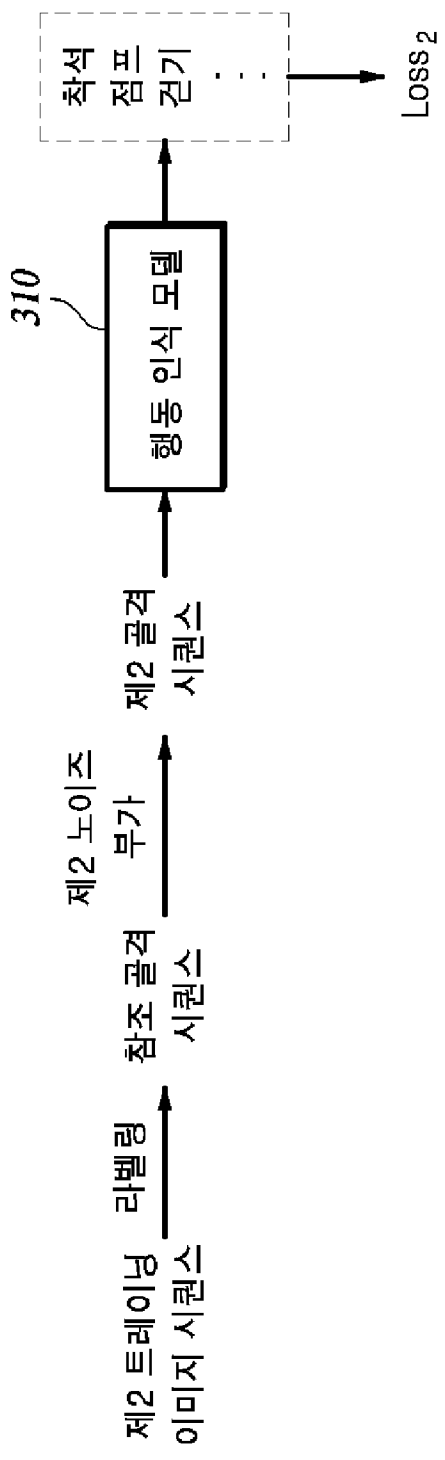
[도 1]



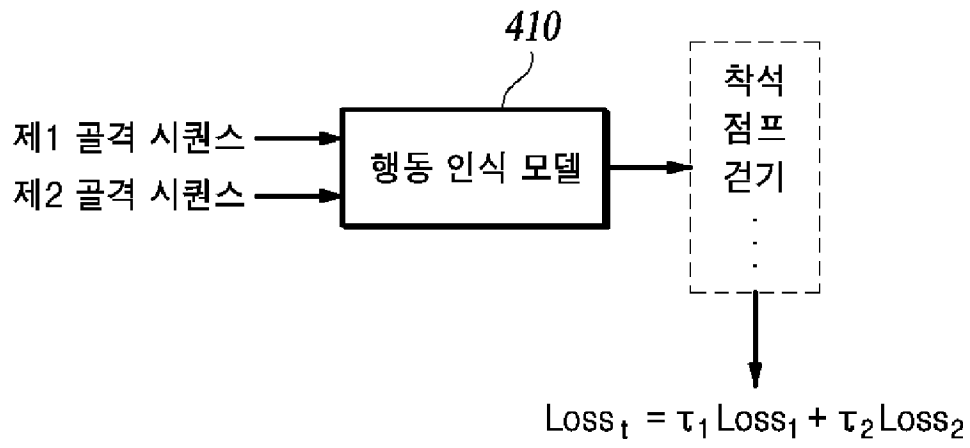
[도2]



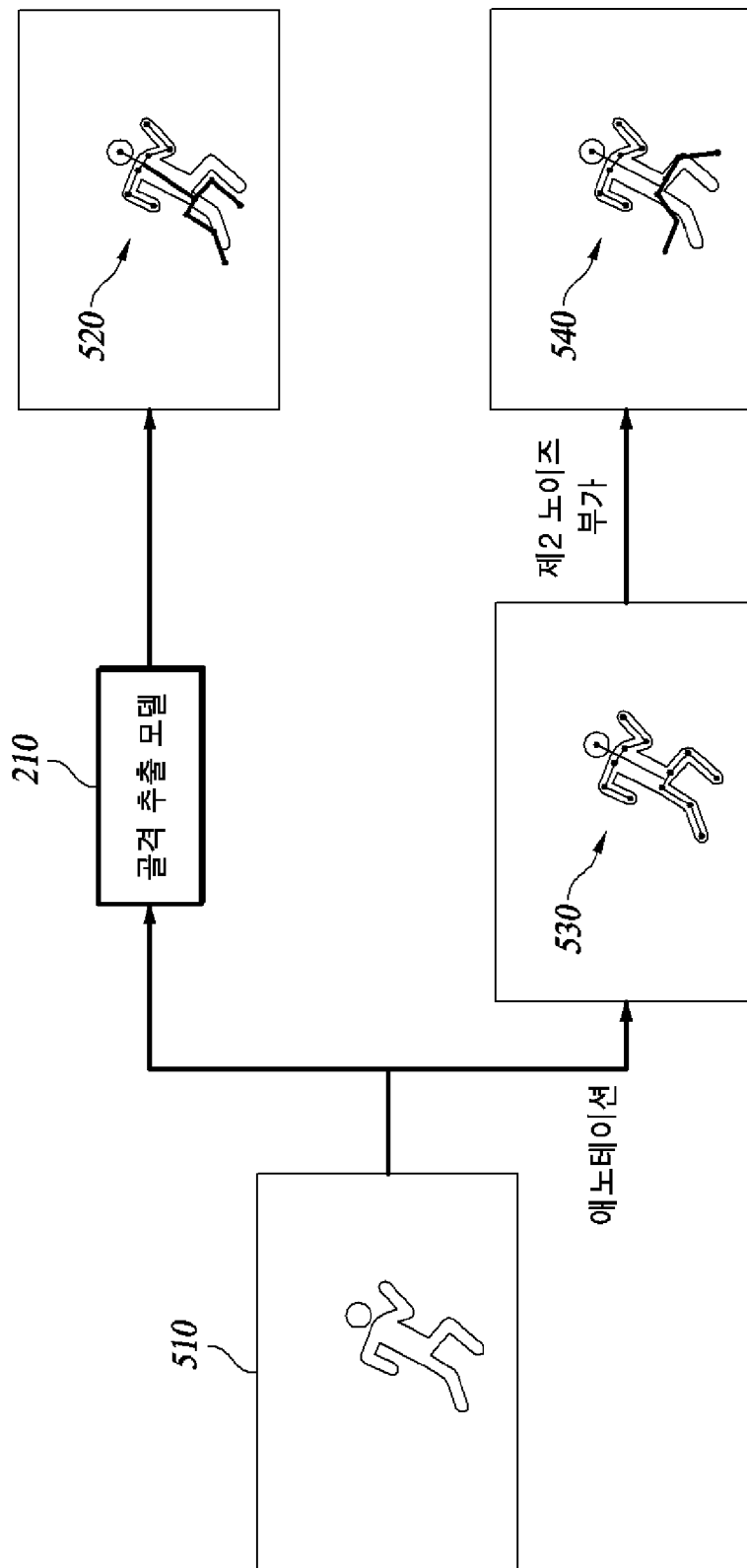
[도3]



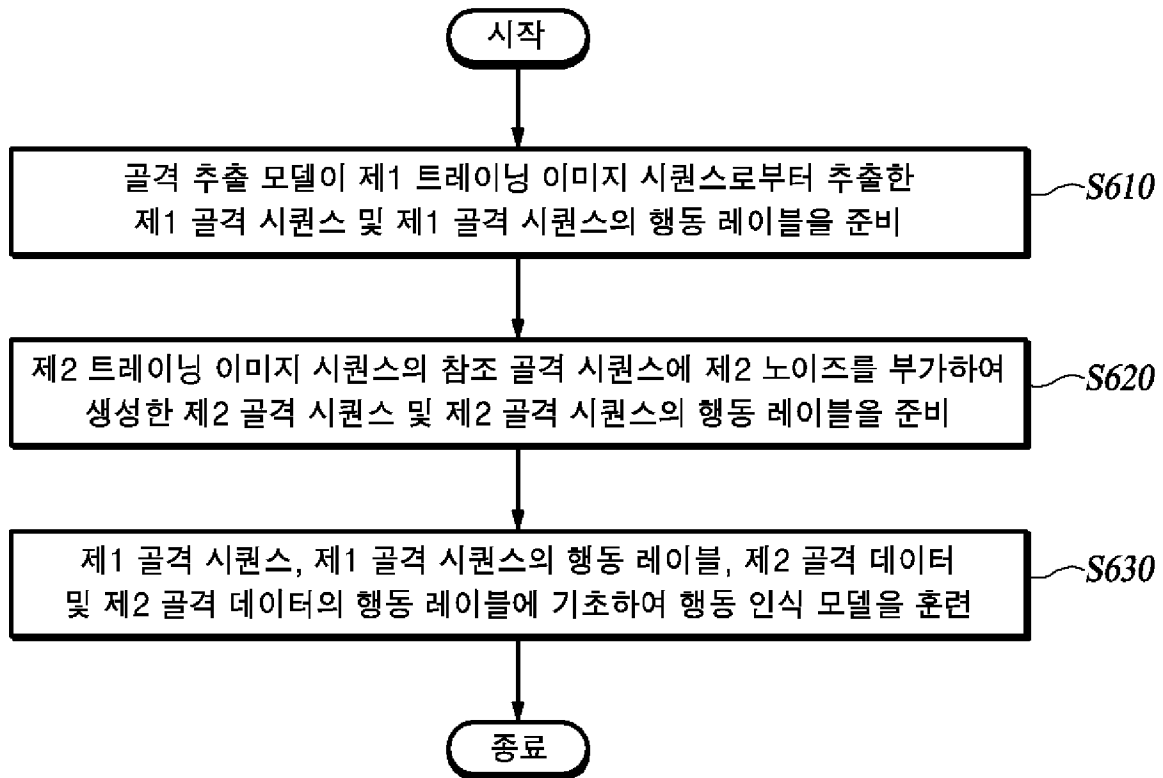
[도4]



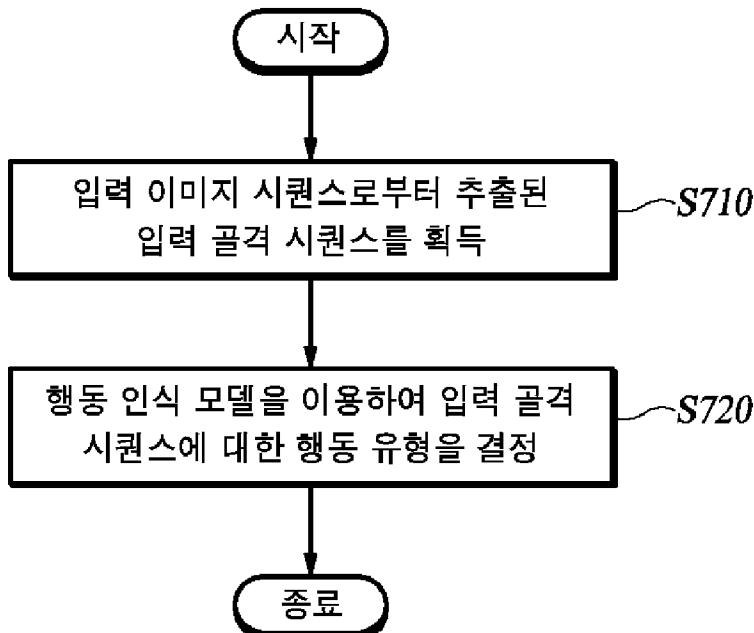
[도5]



[도6]



[도7]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2024/007350

A. CLASSIFICATION OF SUBJECT MATTER G06V 40/20(2022.01)i; G06V 10/778(2022.01)i; G06V 10/82(2022.01)i; G06N 3/04(2006.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06V 40/20(2022.01); A61B 5/11(2006.01); G06F 18/00(2023.01); G06K 9/00(2006.01); G06K 9/48(2006.01); G06K 9/62(2006.01); H04N 21/44(2011.01) Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models: IPC as above Japanese utility models and applications for utility models: IPC as above Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS (KIPO internal) & keywords: 행동 인식(action recognition), 모델(model), 훈련(training), 골격(skeleton), 노이즈(noise)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 105373810 A (TSINGHUA UNIVERSITY) 02 March 2016 (2016-03-02) See paragraph [0056]; claim 1; and figure 1.	1-8
Y	KR 10-2020-0022788 A (KOREA ELECTRONICS TECHNOLOGY INSTITUTE) 04 March 2020 (2020-03-04) See paragraph [0032]; and figure 2.	1-8
A	CN 108875708 A (GUANGDONG UNIVERSITY OF TECHNOLOGY) 23 November 2018 (2018-11-23) See paragraphs [0051]-[0059]; and figure 1.	1-8
A	KR 10-2022-0066535 A (PLASK CO., LTD.) 24 May 2022 (2022-05-24) See paragraphs [0075]-[0090], [0092]-[0093] and [0095]-[0096]; and figure 3.	1-8
A	KR 10-2021-0031405 A (NAVER CORPORATION) 19 March 2021 (2021-03-19) See paragraphs [0067]-[0074]; and figure 3.	1-8
☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents: “A” document defining the general state of the art which is not considered to be of particular relevance “D” document cited by the applicant in the international application “E” earlier application or patent but published on or after the international filing date “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) “O” document referring to an oral disclosure, use, exhibition or other means “P” document published prior to the international filing date but later than the priority date claimed “T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art “&” document member of the same patent family		
Date of the actual completion of the international search 12 September 2024		Date of mailing of the international search report 12 September 2024
Name and mailing address of the ISA/KR Korean Intellectual Property Office Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208 Facsimile No. +82-42-481-8578		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2024/007350

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	105373810	A	02 March 2016	CN	105373810	B	17 January 2020
KR	10-2020-0022788	A	04 March 2020	None			
CN	108875708	A	23 November 2018	None			
KR	10-2022-0066535	A	24 May 2022	KR	10-2023-0093191	A	27 June 2023
				KR	10-2542878	B1	14 June 2023
KR	10-2021-0031405	A	19 March 2021	EP	3792821	A1	17 March 2021
				JP	2021-043974	A	18 March 2021
				US	11625953	B2	11 April 2023
				US	2021-0073525	A1	11 March 2021
				US	2023-0215160	A1	06 July 2023

A. 발명이 속하는 기술분류(국제특허분류(IPC)) G06V 40/20(2022.01)i; G06V 10/778(2022.01)i; G06V 10/82(2022.01)i; G06N 3/04(2006.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) G06V 40/20(2022.01); A61B 5/11(2006.01); G06F 18/00(2023.01); G06K 9/00(2006.01); G06K 9/48(2006.01); G06K 9/62(2006.01); H04N 21/44(2011.01) 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 행동 인식(action recognition), 모델(model), 훈련(training), 골격(skeleton), 노이즈(noise)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	CN 105373810 A (TSINGHUA UNIVERSITY) 2016.03.02 단락 [0056]; 청구항 1; 및 도면 1	1-8
Y	KR 10-2020-0022788 A (전자부품연구원) 2020.03.04 단락 [0032]; 및 도면 2	1-8
A	CN 108875708 A (GUANGDONG UNIVERSITY OF TECHNOLOGY) 2018.11.23 단락 [0051]-[0059]; 및 도면 1	1-8
A	KR 10-2022-0066535 A (주식회사 플라스크) 2022.05.24 단락 [0075]-[0090], [0092]-[0093], [0095]-[0096]; 및 도면 3	1-8
A	KR 10-2021-0031405 A (네이버 주식회사) 2021.03.19 단락 [0067]-[0074]; 및 도면 3	1-8
☐ 추가 문헌이 C(계속)에 기재되어 있습니다. ☒ 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “D” 본 국제출원에서 출원인이 인용한 문헌 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. “&” 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일 2024년09월12일(12.09.2024)		국제조사보고서 발송일 2024년09월12일(12.09.2024)
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578		심사관 정종한 전화번호 +82-42-481-5642

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
CN 105373810 A	2016/03/02	CN 105373810 B	2020/01/17
KR 10-2020-0022788 A	2020/03/04	없음	
CN 108875708 A	2018/11/23	없음	
KR 10-2022-0066535 A	2022/05/24	KR 10-2023-0093191 A	2023/06/27
		KR 10-2542878 B1	2023/06/14
KR 10-2021-0031405 A	2021/03/19	EP 3792821 A1	2021/03/17
		JP 2021-043974 A	2021/03/18
		US 11625953 B2	2023/04/11
		US 2021-0073525 A1	2021/03/11
		US 2023-0215160 A1	2023/07/06