



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0123823
(43) 공개일자 2021년10월14일

(51) 국제특허분류(Int. Cl.)
G06T 7/246 (2017.01) G06N 3/08 (2006.01)
(52) CPC특허분류
G06T 7/251 (2017.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2020-0041501
(22) 출원일자 2020년04월06일
심사청구일자 2020년04월06일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대 학교)
(72) 발명자
이상철
인천광역시 연수구 아트센터대로97번길 76, 1402 동 805호(송도동, 송도 더샵 하버뷰 14단지)
이종민
서울특별시 동작구 신대방길 42, 305호(신대방동)
(74) 대리인
양성보

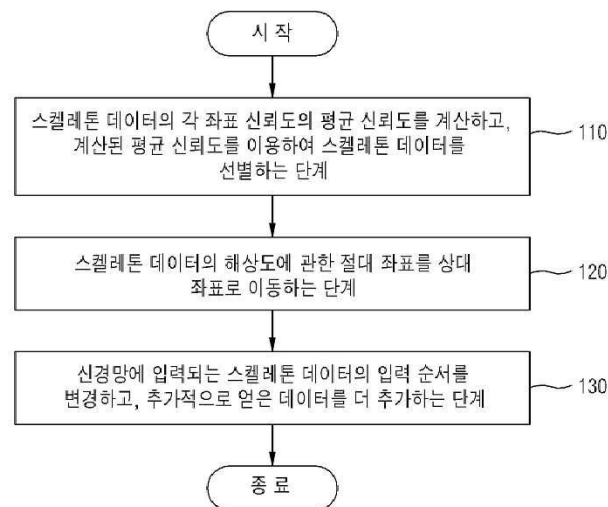
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법 및 장치

(57) 요약

딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법 및 장치가 제시된다. 본 발명에서 제안하는 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법은, 스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별하는 단계, 선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동하는 단계 및 신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가하는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06T 2207/20044 (2013.01)

G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711099481
과제번호	2017M3C8A8091768
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	기초 . 원천기술 개발사업
연구과제명	[Ezbaro] 디지털 컴패니언 기반의 고령자 모니터링 및 대응 소프트웨어 기술 개발
기 여 율	1/2
과제수행기관명	인하대학교
연구기간	2019.09.20 ~ 2020.07.19

이 발명을 지원한 국가연구개발사업

과제고유번호	1711089183
과제번호	2018R1A2B6001643
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구
연구과제명	[Ezbaro] 영상 정합, 추적 및 검출을 위한 기계학습 기반 목적 지향형 특징벡터 합
성 기술	
기 여 율	1/2
과제수행기관명	인하대학교
연구기간	2019.03.01 ~ 2020.02.29

명세서

청구범위

청구항 1

스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별하는 단계;

선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동하는 단계; 및

신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가하는 단계를 포함하는 스켈레톤 데이터의 전 처리 방법.

청구항 2

제1항에 있어서,

스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별하는 단계는,

계산된 평균 신뢰도와 미리 정해진 임계값을 비교하여 임계값 이상의 평균 신뢰도를 갖는 스켈레톤 데이터를 선별하는

스켈레톤 데이터의 전 처리 방법.

청구항 3

제1항에 있어서,

선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동하는 단계는,

신경망을 통해 학습하기 위한 학습 데이터의 양을 감소시키기 위해 스켈레톤 데이터의 미리 정해진 중심 좌표를 기준으로 하여 해상도에 관한 좌표계 이동을 수행하는

스켈레톤 데이터의 전 처리 방법.

청구항 4

신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가하는 단계는,

스켈레톤 데이터의 상대 좌표로의 이동에 대한 정보의 손실을 방지하기 위해 이전 데이터와 현재 데이터의 차이를 구하여 신경망의 입력에 추가하고, 상대 좌표 데이터와 추가한 데이터의 입력 순서를 바꾸어 신경망에 입력하는

스켈레톤 데이터의 전 처리 방법.

청구항 5

스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별하는 신뢰도 선별부;

선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동하는 좌표 이동부; 및

신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가하는 정보 추가부

를 포함하는 스켈레톤 데이터의 전 처리 장치.

청구항 6

제5항에 있어서,

신뢰도 선별부는,

신경망을 통해 학습하기 위한 학습 데이터의 양을 감소시키기 위해 스켈레톤 데이터의 미리 정해진 중심 좌표를 기준으로 하여 해상도에 관한 좌표계 이동을 수행하는

스켈레톤 데이터의 전 처리 장치.

청구항 7

제5항에 있어서,

좌표 이동부는,

신경망을 통해 학습하기 위한 학습 데이터의 양을 감소시키기 위해 스켈레톤 데이터의 미리 정해진 중심 좌표를 기준으로 하여 해상도에 관한 좌표계 이동을 수행하는

스켈레톤 데이터의 전 처리 장치.

청구항 8

제5항에 있어서,

정보 추가부는,

스켈레톤 데이터의 상대 좌표로의 이동에 대한 정보의 손실을 방지하기 위해 이전 데이터와 현재 데이터의 차이를 구하여 신경망의 입력에 추가하고, 상대 좌표 데이터와 추가한 데이터의 입력 순서를 바꾸어 신경망에 입력하는

스켈레톤 데이터의 전 처리 장치.

발명의 설명

기술 분야

[0001] 본 발명은 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법 및 장치에 관한 것이다.

배경 기술

[0002] 인간의 행동 인식은 여러 분야에서 활발한 연구가 진행되고 있다. 행동 인식의 성능은 특징의 추출과 그 정보를 얼마나 잘 사용하는지에 좌우된다.

[0003] 컨볼루션(Convolutional) 신경망은 이미지에서 뛰어난 특징 정보의 추출이 가능하지만 단일 이미지를 염두에 둔 신경망이기 때문에 시간적 정보를 잘 추출하지 못하며 이미지만 입력으로 받을 수 있는 단점이 있다.

[0004] 순환 신경망(Recurrent Neural Network)은 시간적 순서를 가지는 데이터에 특화된 신경망이다. 하지만 이미지를 입력으로 받기 어렵고 긴 시간적 데이터에 대해서는 좋지 못한 성능을 보이는 경향이 있다.

[0005] LSTM(Long Short-Term Memory) 신경망은 순환 신경망의 긴 시간적 데이터를 처리하지 못하는 단점을 보완한 신경망이다. 이전 데이터로부터 필요한 정보만을 추출하는 것으로 긴 시간적 데이터의 문제를 어느 정도 해결하였다.

[0006] 일반적인 행동 인식 네트워크는 특징 정보를 잘 추출하는 컨볼루션 신경망과 시간적 정보를 잘 파악하는 LSTM 신경망을 같이 사용한다. 이러한 방법들은 좋은 성능을 보장하지만 무거운 네트워크와 큰 용량을 가지는 데이터로 인해 실생활에 적용하기 어렵다. 행동 인식 네트워크를 실생활에 적용하기 위해서는 작은 용량을 가지는 데이터와 가벼운 네트워크가 요구된다.

발명의 내용

해결하려는 과제

[0007] 본 발명이 이루고자 하는 기술적 과제는 특정 신뢰도 이상의 데이터만 선별해 사용하는 신뢰도 선별 방법과 스

스켈레톤의 절대 좌표를 상대 좌표로 이동하는 좌표 이동 방법 및 네트워크에 들어가는 데이터의 입력 순서의 변경과 추가적으로 얻은 데이터를 더 넣어주는 정보 추가 방법을 이용하는 전 처리 방법 및 장치를 제공하는데 있다.

과제의 해결 수단

- [0008] 일 측면에 있어서, 본 발명에서 제안하는 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법은, 스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별하는 단계, 선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동하는 단계 및 신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가하는 단계를 포함한다.
- [0009] 스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별하는 단계는 계산된 평균 신뢰도와 미리 정해진 임계값을 비교하여 임계값 이상의 평균 신뢰도를 갖는 스켈레톤 데이터를 선별한다.
- [0010] 선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동하는 단계는 신경망을 통해 학습하기 위한 학습 데이터의 양을 감소시키기 위해 스켈레톤 데이터의 미리 정해진 중심 좌표를 기준으로 하여 해상도에 관한 좌표계 이동을 수행한다.
- [0011] 신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가하는 단계는 스켈레톤 데이터의 상대 좌표로의 이동에 대한 정보의 손실을 방지하기 위해 이전 데이터와 현재 데이터의 차이를 구하여 신경망의 입력에 추가하고, 상대 좌표 데이터와 추가한 데이터의 입력 순서를 바꾸어 신경망에 입력한다.
- [0012] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 장치는 스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별하는 신뢰도 선별부, 선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동하는 좌표 이동부 및 신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가하는 정보 추가부를 포함한다.

발명의 효과

- [0013] 본 발명의 실시예들에 따르면 특정 신뢰도 이상의 데이터만 선별해 사용하는 신뢰도 선별 방법과 스켈레톤의 절대 좌표를 상대 좌표로 이동하는 좌표 이동 방법 및 네트워크에 들어가는 데이터의 입력 순서의 변경과 추가적으로 얻은 데이터를 더 넣어주는 정보 추가 방법을 이용하여 더욱 효과적인 전 처리를 수행할 수 있다.

도면의 간단한 설명

- [0014] 도 1은 본 발명의 일 실시예에 따른 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법을 설명하기 위한 흐름도이다.
- 도 2는 본 발명의 일 실시예에 따른 좌표 이동 과정을 설명하기 위한 도면이다.
- 도 3은 본 발명의 일 실시예에 따른 LSTM 신경망 구조를 나타내는 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 스켈레톤 데이터 구조를 나타내는 도면이다.
- 도 5는 본 발명의 일 실시예에 따른 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 장치의 구성을 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0015] 본 발명에서는 LSTM을 이용한 2차원 스켈레톤 데이터의 행동 인식 정확도 향상을 위한 다양한 데이터 전 처리 기법을 제안한다. 본 발명의 실시예에 따르면, 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법 및 장치를 통해 가벼운 네트워크인 기본 LSTM 신경망과 2차원 스켈레톤 데이터로 여러 전 처리 기법을 사용하여 실험을 진행하고 전 처리 기법이 성능에 미치는 결과를 비교하여 더욱 효과적인 전 처리 기법을 제시한다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

- [0017] 도 1은 본 발명의 일 실시예에 따른 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법을 설명하기 위한 흐름도이다.
- [0018] 제안하는 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법은 스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별하는 단계(110), 선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동하는 단계(120) 및 신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가하는 단계(130)를 포함한다.
- [0019] 단계(110)에서, 스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별한다. 계산된 평균 신뢰도와 미리 정해진 임계값을 비교하여 임계값 이상의 평균 신뢰도를 갖는 스켈레톤 데이터를 선별한다. 신뢰성 높은 데이터가 가져오는 성능적 이점을 보기 위해 특정 신뢰도 이상의 데이터만 선별해 사용하는 신뢰도 선별 방법을 이용한다.
- [0020] 일반적인 스켈레톤 데이터는 좌표의 값과 각 좌표에 대한 신뢰도를 포함한다. 신뢰도 선별 방법은 스켈레톤 데이터의 각 좌표 신뢰도의 평균을 계산한다. 이렇게 계산된 평균 신뢰도와 미리 정해진 임계값을 비교하여 평균 신뢰도가 임계값을 넘기지 못하면 데이터를 무시하고 다음 데이터를 판단한다. 본 발명의 일 실시예에 따르면, 임계값을 평균 신뢰도 50% 이상으로 설정하여 실험을 진행하였다.
- [0021] 단계(120)에서, 선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동한다. 신경망을 통해 학습하기 위한 학습 데이터의 양을 감소시키기 위해 스켈레톤 데이터의 미리 정해진 중심 좌표를 기준으로 하여 해상도에 관한 좌표계 이동을 수행한다.
- [0023] 도 2는 본 발명의 일 실시예에 따른 좌표 이동 과정을 설명하기 위한 도면이다.
- [0024] 도 2(a)는 절대 좌표계에서 여러 공간 상에서 같은 동작을 취하는 상태를 나타내는 도면이다. 도 2(b)는 미리 정해진 중심 좌표(0, 0)를 기준으로 하는 상대 좌표계에서의 동작을 취하는 상태를 나타내는 도면이다.
- [0025] 해상도를 좌표로 가지는 절대 좌표의 스켈레톤 데이터는 도 2(a)와 같이 영상 내에 사람의 위치가 달라지면 같은 동작이라도 좌표 값의 차이가 발생한다. LSTM 신경망이 이러한 절대 좌표의 스켈레톤 데이터를 학습하려면 영상 내에 사람이 존재 할 수 있는 모든 위치에 대한 데이터를 필요로 하게 되고 그에 따라 학습 시켜야 할 데이터의 양이 증가하게 된다.
- [0026] 이와 반대로, 도 2(b)와 같이 상대 좌표를 가지는 스켈레톤 데이터는 중심 좌표(0, 0)에 고정되어 행동을 취하기 때문에 해상도 위치에 따른 종속성이 사라지게 된다. 따라서 절대 좌표에 비해 많은 학습 데이터를 필요로 하지 않게 되고 그에 따라 성능 또한 크게 증가할 수 있다.
- [0027] 상대 좌표를 가지는 스켈레톤 데이터는 스켈레톤의 특정 좌표(0,0)를 기준으로 하여 좌표계 이동을 수행한다. 일반적으로, 골반의 중앙 좌표나 목의 좌표를 중심 좌표로 선택하여 상대 좌표로 좌표계 이동을 하지만, 본 발명의 일 실시예에 따르면, 사용하는 스켈레톤 데이터에는 골반의 중앙 좌표와 목의 좌표 값이 없기 때문에 코의 좌표를 중심 좌표로 선택하여 상대 좌표로 좌표계 이동을 시행하였다.
- [0028] 단계(130)에서, 신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가한다. 스켈레톤 데이터의 상대 좌표로의 이동에 대한 정보의 손실을 방지하기 위해 이전 데이터와 현재 데이터의 차이를 구하여 신경망의 입력에 추가하고, 상대 좌표 데이터와 추가한 데이터의 입력 순서를 바꾸어 신경망에 입력한다.
- [0029] 스켈레톤 데이터를 상대 좌표로 이동하게 되면 정확도는 증가하지만 그와 반대로 좌표계가 고정되어 영상 내 이동에 대한 정보의 손실을 가져온다. 신경망에 들어가는 데이터의 입력 순서의 변경과 추가적으로 얻은 데이터를 더 넣어주는 정보 추가 방법은 이러한 이동에 대한 정보의 손실을 방지하고자 이전 데이터와 현재 데이터의 차이를 구하여 명시적으로 LSTM 신경망의 입력에 추가하였다. 추가적으로, 좌표데이터와 명시적으로 추가한 데이터의 입력 순서를 바꾸어 실험하면서 입력 순서가 LSTM 신경망 결과에 영향이 있는 실험을 진행하였다.

[0031] 도 3은 본 발명의 일 실시예에 따른 LSTM 신경망 구조를 나타내는 도면이다.

[0032] 순환 신경망의 대표적 종류는 RNN, LSTM, GRU(Gated Recurrent Unit)가 있다. RNN은 순환 신경망의 가장 기본적인 구조를 가지고 있다. 짧은 시간적 데이터를 잘 처리하지만 긴 시간적 데이터는 잘 처리하지 못한다. GRU는 네트워크가 가볍다는 장점이 있지만 그만큼 성능이 좋지 못하다. 따라서, 본 발명의 실시예에 따르면, 긴 시간적 데이터를 비교적 잘 처리하는 LSTM 신경망을 실험에 사용하였다.

[0033] LSTM 신경망의 구조는 기본적으로 전개 크기, 신경망 셀과 레이어와 같이 RNN과 같은 구조를 가진다. 본 발명에서는 전 처리 실험마다 레이어와 히든 출력의 크기를 바꾸어 가면서 최적의 성능을 찾아 정확도를 비교하였다.

[0034] LSTM 신경망의 하이퍼 파라미터는 전개 크기, 입력 크기, 히든 출력의 크기, 레이어의 수 등이 있다. 본 발명의 실시예에 따르면, 전개 크기는 8로 고정하였고 입력 크기는 전 처리 기법에 따라 달라진다.

[0035] 본 발명에서는 각 실험마다 아래의 표 1과 같이 최적의 LSTM 신경망의 하이퍼 파라미터를 찾기 위한 실험을 진행하였다. 표 1은 LSTM 신경망 하이퍼 파라미터 설정 실험을 위한 데이터를 나타낸다.

[0036] <표 1>

레이어의 수	1, 2, 3, 4, 5, 6
히든 출력의 크기	16, 32, 64, 128, 256, 512

[0037]

[0039] 도 4는 본 발명의 일 실시예에 따른 스켈레톤 데이터 구조를 나타내는 도면이다.

[0040] 본 발명의 실시예에 따르면, 스켈레톤 데이터는 도 4와 같이 17개의 좌표를 가진다. 각 좌표는 순서대로 코, 왼쪽 눈, 오른쪽 눈, 왼쪽 귀, 오른쪽 귀, 왼쪽 어깨, 오른쪽 어깨, 왼쪽 팔꿈치, 오른쪽 팔꿈치, 왼쪽 손목, 오른쪽 손목, 왼쪽 엉덩이, 오른쪽 엉덩이, 왼쪽 무릎, 오른쪽 무릎, 왼쪽 발목, 오른쪽 발목이며 640×480 해상도의 절대 좌표로 각 좌표의 신뢰도와 함께 기록된다.

[0041] 스켈레톤 추출을 위한 영상 데이터는 자체적으로 취득한 영상 데이터를 사용하였다. 스켈레톤 데이터의 추출 또한, 오픈 포즈가 아닌 별도로 개발된 스켈레톤 추출 알고리즘을 사용하였다. 영상 데이터는 초당 8장의 이미지로 변환하여 이미지에서 스켈레톤 데이터를 추출하였다.

[0042] 총 19명을 촬영 하였으며 학습용 스켈레톤 데이터는 4명의 촬영 영상(693개)을 사용하였고 테스트용 스켈레톤 데이터는 15명의 촬영 영상(497개)을 사용하였다. 영상 취득은 일상적인 11가지의 행동을 취득하였다. 표 2는 취득한 행동의 종류를 나타낸다.

[0043] <표 2>

행동	
바닥에 앉아있다 눕는 동작	바닥에 앉아있다 서있는 동작
바닥에 앉아있는 동작	누워있다 바닥에 앉는 동작
누워있는 동작	의자에 앉아있는 동작
서있다가 바닥에 앉는 동작	서있다가 의자에 앉는 동작
서있는 동작	걷는 동작
의자에 앉아있다 일어서는 동작	

[0044]

[0046] 도 5는 본 발명의 일 실시예에 따른 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 장치의 구성을 나타내는 도면이다.

[0047] 제안하는 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 장치(500)는 신뢰도 선별부(510), 좌표 이동부(520) 및 정보 추가부(530)를 포함한다.

[0048] 신뢰도 선별부(510)는 스켈레톤 데이터의 각 좌표 신뢰도의 평균 신뢰도를 계산하고, 계산된 평균 신뢰도를 이용하여 스켈레톤 데이터를 선별한다. 계산된 평균 신뢰도와 미리 정해진 임계값을 비교하여 임계값 이상의 평균 신뢰도를 갖는 스켈레톤 데이터를 선별한다. 신뢰성 높은 데이터가 가져오는 성능적 이점을 보기 위해 특정 신

되도 이상의 데이터만 선별해 사용하는 신뢰도 선별 방법을 이용한다.

[0049] 일반적인 스켈레톤 데이터는 좌표의 값과 각 좌표에 대한 신뢰도를 포함한다. 신뢰도 선별 방법은 스켈레톤 데이터의 각 좌표 신뢰도의 평균을 계산한다. 이렇게 계산된 평균 신뢰도와 미리 정해진 임계값을 비교하여 평균 신뢰도가 임계값을 넘기지 못하면 데이터를 무시하고 다음 데이터를 판단한다. 본 발명의 일 실시예에 따르면, 임계값을 평균 신뢰도 50% 이상으로 설정하여 실험을 진행하였다.

[0050] 좌표 이동부(520)는 선별된 스켈레톤 데이터의 해상도에 관한 절대 좌표를 상대 좌표로 이동한다. 신경망을 통해 학습하기 위한 학습 데이터의 양을 감소시키기 위해 스켈레톤 데이터의 미리 정해진 중심 좌표를 기준으로 하여 해상도에 관한 좌표계 이동을 수행한다.

[0051] 정보 추가부(530)는 신경망에 입력되는 스켈레톤 데이터의 입력 순서를 변경하고, 추가적으로 얻은 데이터를 더 추가한다. 스켈레톤 데이터의 상대 좌표로의 이동에 대한 정보의 손실을 방지하기 위해 이전 데이터와 현재 데이터의 차이를 구하여 신경망의 입력에 추가하고, 상대 좌표 데이터와 추가한 데이터의 입력 순서를 바꾸어 신경망에 입력한다.

[0052] 스켈레톤 데이터를 상대 좌표로 이동하게 되면 정확도는 증가하지만 그와 반대로 좌표계가 고정되어 영상 내 이동에 대한 정보의 손실을 가져온다. 신경망에 들어가는 데이터의 입력 순서의 변경과 추가적으로 얻은 데이터를 더 넣어주는 정보 추가 방법은 이러한 이동에 대한 정보의 손실을 방지하고자 이전 데이터와 현재 데이터의 차이를 구하여 명시적으로 LSTM 신경망의 입력에 추가하였다. 추가적으로, 좌표데이터와 명시적으로 추가한 데이터의 입력 순서를 바꾸어 실험하면서 입력 순서가 LSTM 신경망 결과에 영향이 있는 실험을 진행하였다.

[0054] 본 발명의 실시예에 따른 딥러닝 기반 행동 인식 정확도 향상을 위한 스켈레톤 데이터의 전 처리 방법을 통한 실험을 진행함에 앞서 모든 실험에 여러 사람이 등장할 경우 판단해야 할 사람을 정하는 사람 선택 기능과 선택한 사람만을 추적하여 우선적으로 판단하는 추적 기능을 기본적으로 적용하였다.

[0055] 사람 선택 기능은 다수의 사람이 나올 경우 각 사람마다 좌표의 전체 신뢰도 평균을 구하여 평균이 가장 높은 사람을 선택하게 구현하였다

[0056] 추적 기능은 선택된 사람이 다음 데이터에 존재할 때 사람 선택 기능을 수행하지 않고 이미 선택된 사람을 우선적으로 판단을 수행할 수 있게 구현하였다.

[0057] 순환 신경망은 이전 시간의 데이터와 일관된 연관성이 중요하기 때문에 전 처리 기법으로 입력 시간 중간에 데이터가 필터링 되면 네트워크에 입력을 하지 않고 다음 시간 데이터부터 전 처리를 다시 수행하여 네트워크의 입력으로 사용하게 구현하였고 표 3과 같이 순환 신경망의 하이퍼 파라미터 실험을 진행하여 손실이 가장 낮은 학습의 정확도로 결과를 비교하였다. 표 3은 하이퍼 파라미터, 히든 출력의 크기에 대한 정확도 비교를 나타낸다. 이때, 레이어는 1이고, 히든 출력의 크기는 128이다.

[0058] <표 3>

기능 \ 히든 크기	16	32	64	128	256	512
절대 좌표	39.34	34.64	44.73	48.84	38.38	43.53
+ 좌표이동	69.51	69.28	71.37	69.32	72.05	73.10
+ 신뢰도	68.46	67.55	75.62	76.46	77.05	72.64
+ 정보추가	66.46	71.94	74.11	78.43	75.65	75.19
기능 \ 레이어 크기	1	2	3	4	5	6
절대 좌표	48.84	48.51	36.41	37.75	25.88	29.43
+ 좌표이동	69.32	68.00	66.08	66.41	64.84	62.21
+ 신뢰도	76.46	73.11	68.47	70.42	67.25	56.98
+ 정보추가	78.43	68.15	70.68	67.22	65.64	57.52

[0059]

[0060] 먼저 일반적으로 많이 쓰이는 전 처리 방법인 좌표 이동 방법의 효과를 검증하기 위해 절대 좌표로 학습시킨 LSTM 신경망과 상대 좌표로 학습시킨 LSTM 신경망을 비교하였다. 절대 좌표로 학습한 LSTM의 경우 정확도는 48.8%, 손실은 1.62가 나왔다. 상대 좌표로 학습한 LSTM의 경우 정확도는 73.1%, 손실은 0.82가 나왔다. 좌표 이동 방법을 적용했을 때 정확도는 24.3% 증가하고, 손실은 0.8 감소하여 뛰어난 효과를 입증하였다.

[0061] 여기에서 신뢰도 선별 방법을 적용하게 되면 정확도는 77.0%, 손실은 0.72로 좌표 이동 방법에 비해 정확도는 3.9% 증가하고, 손실은 0.1감소하여 성능 향상에 도움이 되었다.

[0062] 이에 정보 추가 방법을 적용하게 되면 정확도는 78.4%, 손실은 0.67로 정확도는 1.4% 증가하고 손실은 0.05 감소하였다. 정확도의 증가와 손실의 감소 폭이 상대적으로 적지만 이동에 대한 정보를 크게 가지고 있는 움직임이 있는 행동에 대해서 평균적으로 20% 이상의 정확도 향상을 보였다.

[0063] 표 4는 입력 순서에 대한 정확도 비교를 나타낸다. 이때, 레이어는 1이고, 히든 출력의 크기는 128이다.

[0064] <표 4>

입력 순서	상대좌표 x_t 이전 데이터 차이 x_t , 상대좌표 y_t 이전 데이터 차이 y_t	상대좌표 (x_t, y_t) , 이전 데이터 차이 (x_t, y_t)
정확도	78.43	75.01

[0065]

[0066] 표 5는 정보 추가에 대한 행동 별 정확도 비교를 나타낸다.

[0067] <표 5>

행동	+신뢰도	+정보추가	차이
바닥에 앉아있다 듣는 동작	64	74	+10
바닥에 앉아있다 서있는 동작	31	33	+2
바닥에 앉아있는 동작	71	56	-15
누워있다 바닥에 앉는 동작	3	36	+30
누워있는 동작	15	43	+28
의자에 앉아있다 일어서는 동작	69	88	+17
의자에 앉아있는 동작	96	94	-2
서있다가 바닥에 앉는 동작	36	40	+4
서있다가 의자에 앉는 동작	39	89	+50
서있는 동작	79	82	+3
걷는 동작	83	82	-1

[0068]

[0069] 본 발명에서는 여러 전 처리 기법을 실험하여 그 효과를 입증하였다. 최종적으로 정확도는 29.6% 증가하였고, 일반적인 전 처리와 비교하여 정확도는 5.3% 증가하였다. 또한 이동에 대한 정보가 있는 동작에 대해서 상대적으로 정확도가 증가했음을 확인하였다. 하드웨어 사양이 제한되는 활용에서 전 처리 기법이 어느 정도 성능을 보장할 수 있을 것이다.

[0071] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0072] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로

(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

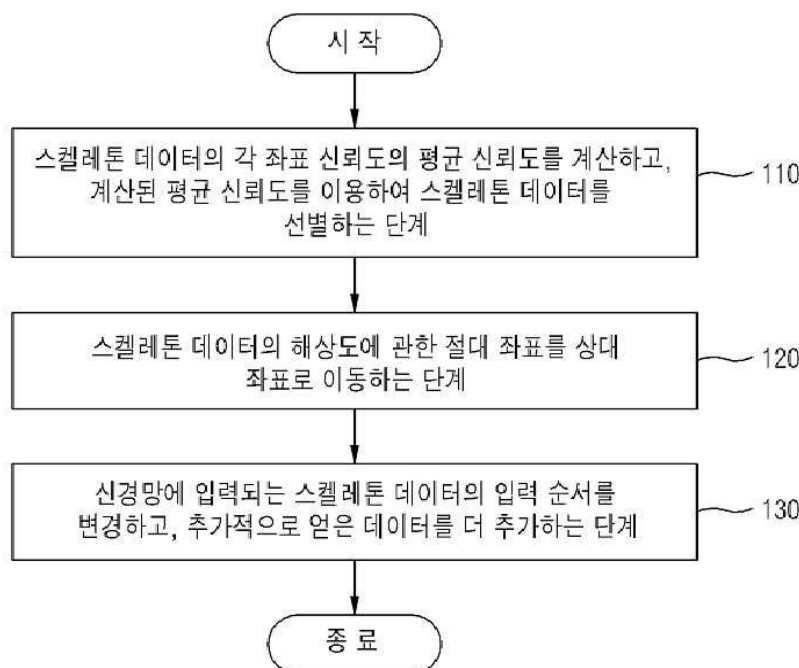
[0073] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0074] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

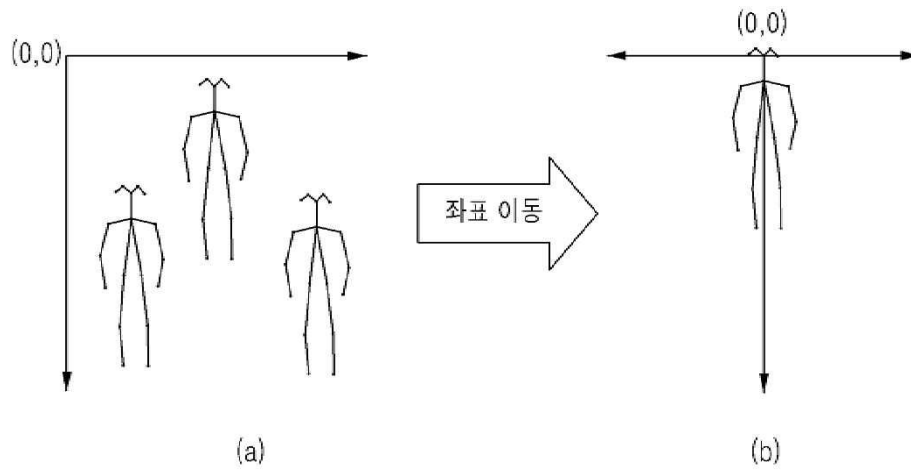
[0075] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

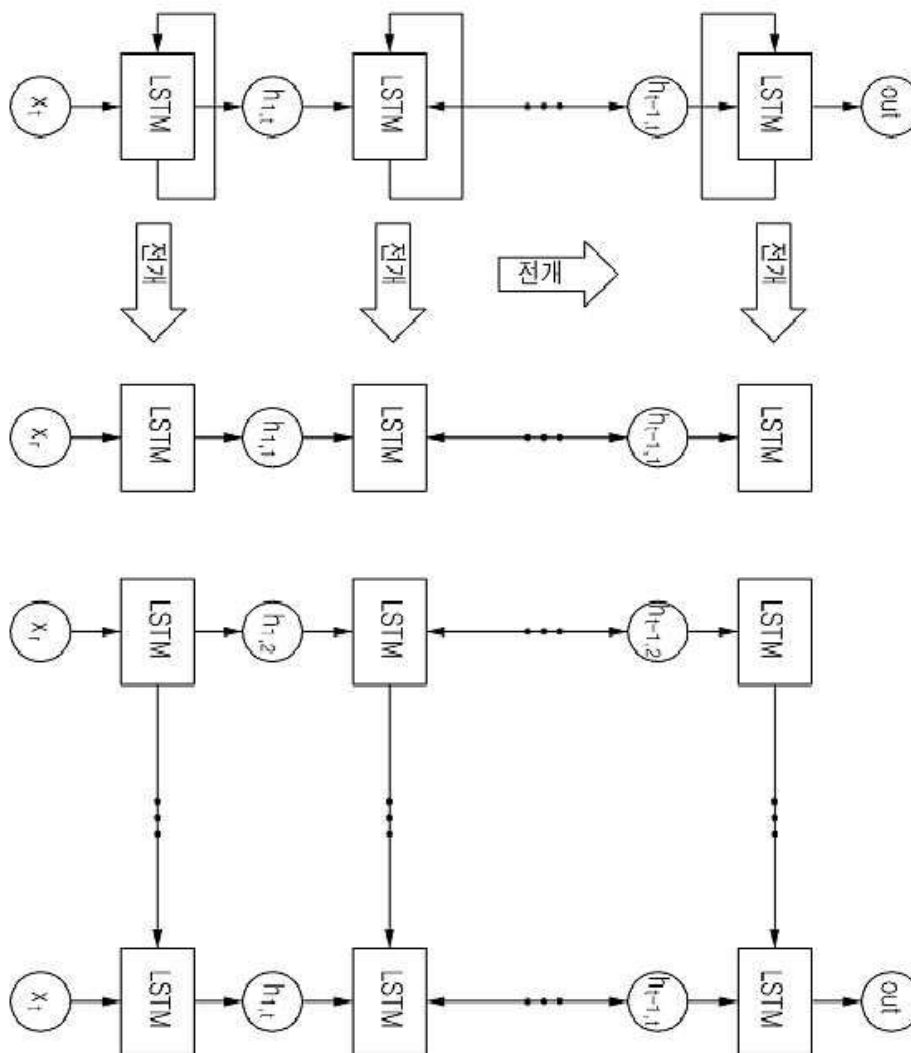
도면1



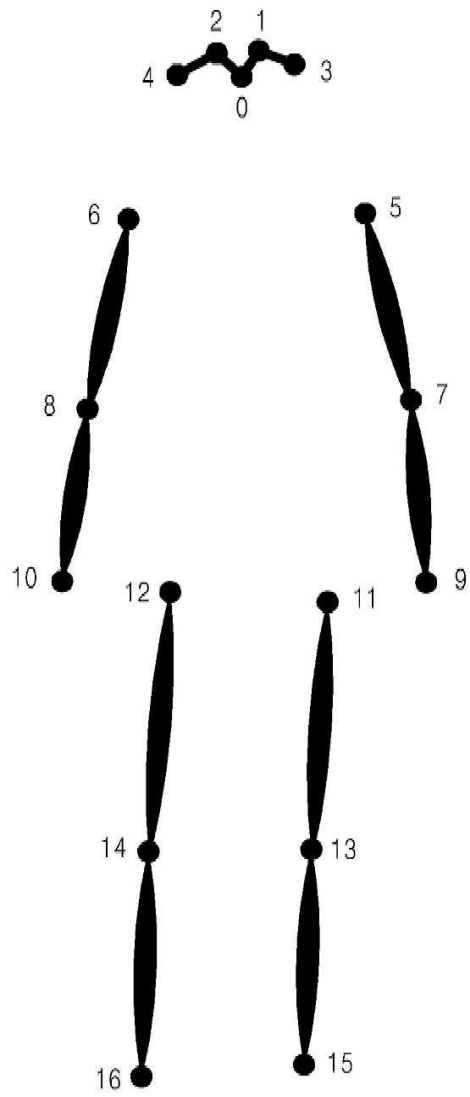
도면2



도면3



도면4



도면5

