(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0136441
(43) 공개일자 2019년12월10일

(51) 국제특허분류(Int. Cl.)
G06F 3/01 (2006.01) *G06K 9/42* (2006.01)
G06N 3/08 (2006.01)
(52) CPC특허분류
G06F 3/017 (2013.01)
G06F 3/011 (2013.01)
(21) 출원번호 10-2018-0062146
(22) 출원일자 2018년05월30일
심사청구일자 2018년05월30일

(71) 출원인
계명대학교 산학협력단
대구광역시 달서구 달구벌대로 1095, 계명대학교 산학협력관 201호(신당동)
(72) 발명자
채지훈
대구광역시 동구 화랑로11길 66 (신천동, 동아아파트)
강수명
대구광역시 북구 칠성남로 54, 702호 (고성동1가, 유성패밀리아)
(뒷면에 계속)
(74) 대리인
김건우

전체 청구항 수 : 총 20 항

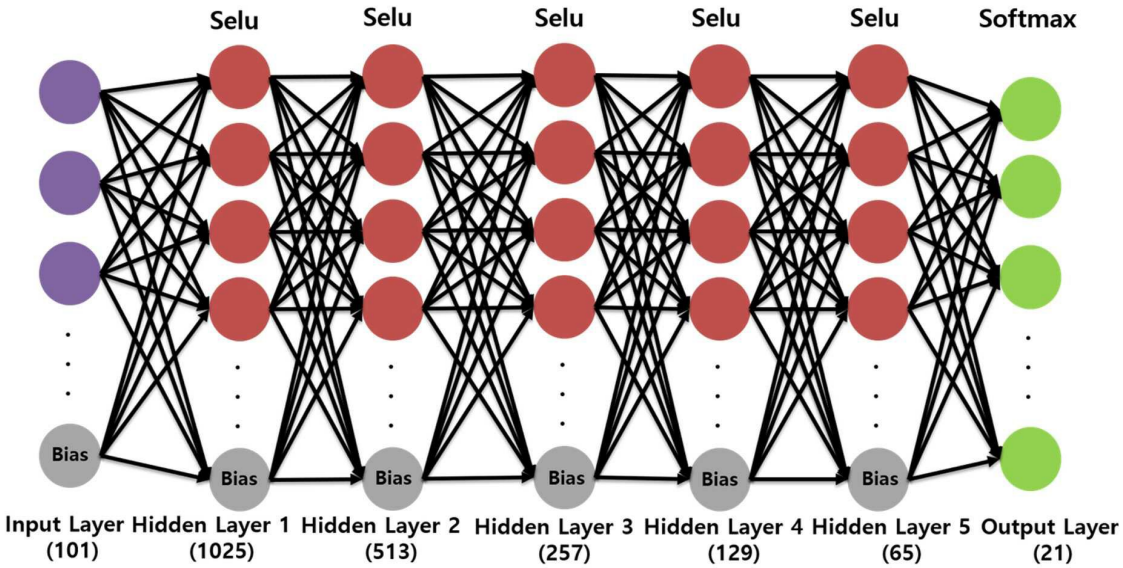
(54) 발명의 명칭 **시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템 및 방법**

(57) 요 약

본 발명은 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템에 관한 것으로서, 보다 구체적으로는 실시간 3차원 제스처 인식 시스템으로서, 모션 입력 장치를 활용하여 실시간으로 입력된 제스처 데이터를 수집하는 제스처 데이터 수집부; 상기 수집된 제스처 데이터에 대해 정규화를 실시하는 정규화부;

(뒷면에 계속)

대 표 도 - 도11



상기 수집된 제스처 데이터와 상기 정규화부에 의해 정규화 된 데이터를 저장하는 데이터 저장부; 및 상기 데이터 저장부에 의해 저장된 데이터를 이용해 딥 러닝(Deep learning)에 기초해 학습하여 딥 러닝 기반의 3차원 제스처 인식 모델을 생성하는 모델 생성부를 포함하는 것을 그 구성상의 특징으로 한다.

또한, 본 발명은 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법에 관한 것으로서, 보다 구체적으로는 실시간 3차원 제스처 인식 방법으로서, (1) 모션 입력 장치를 활용하여 실시간으로 입력된 제스처 데이터를 수집하는 단계; (2) 상기 단계 (1)에 의해 수집된 제스처 데이터에 대해 정규화를 실시하는 단계; (3) 상기 단계 (1)에 의해 수집된 제스처 데이터와 상기 단계 (2)에 의해 정규화 된 데이터를 저장하는 단계; 및 (4) 상기 단계 (3)에 의해 저장된 데이터를 이용해 딥 러닝(Deep Learning)에 기초해 학습하여 딥 러닝 기반의 3차원 제스처 인식 모델을 생성하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

본 발명에서 제안하고 있는 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템 및 방법에 따르면, 실시간 처리가 필요한 계산 효율 및 성능을 얻기 위해 입력 벡터에 대하여 시간적 및 공간적 정규화를 실시하여 데이터의 손실률을 낮추고, 제스처를 그리는 3차원 공간의 정보를 그대로 활용하여 신뢰도와 정확도를 높일 수 있다.

또한, 본 발명에 따르면, 심층 구조를 적용하고 활성화 함수 SELU(Scaled Exponential Linear Unit)를 사용하여, 빠른 학습 성능과 높은 인식률을 갖고, 시간적 및 공간적 정규화 방법을 통해 딥 러닝 기반의 제스처 인식 모델을 학습하며, 더 나아가 소규모의 데이터에서도 빠르게 학습할 수 있다.

(52) CPC특허분류
G06K 9/42 (2013.01)
G06N 3/08 (2013.01)
(72) 발명자
김해성
경상북도 구미시 상사서로 11 우방신세계타운 106동 1802호
이창준
대구광역시 달서구 한들로 36 장기영남네오빌파크 103동 803호
조효준
대구광역시 동구 율하동로12길 7
이준재
대구광역시 수성구 청수로 261 캐슬골드파크4단지 1413-2301호

이 발명을 지원한 국가연구개발사업
과제고유번호 C0512108
부처명 중소벤처기업부
연구관리전문기관 중소기업기술정보진흥원
연구사업명 산학 공동기술개발지원사업
연구과제명 딥 러닝 기반 가상현실 인지재활 게임 콘텐츠 개발
기 여 율 1/1
주관기관 계명대학교 산학협력단
연구기간 2017.06.01 ~ 2018.05.31

명 세 서

청구범위

청구항 1

제스처 인식 시스템으로서,

모션 입력 장치를 활용하여 실시간으로 입력된 제스처 데이터를 수집하는 제스처 데이터 수집부(100);

상기 수집된 제스처 데이터에 대해 정규화를 실시하는 정규화부(200);

상기 수집된 제스처 데이터와 상기 정규화부(200)에 의해 정규화 된 데이터를 저장하는 데이터 저장부(300); 및

상기 데이터 저장부(300)에 의해 저장된 데이터를 이용해 딥 러닝(Deep learning)에 기초해 학습하여 딥 러닝 기반의 3차원 제스처 인식 모델을 생성하는 모델 생성부(400)를 포함하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 2

제1항에 있어서, 상기 제스처 데이터 수집부(100)는,

상기 모션 입력 장치를 활용하고 게임 엔진을 통해서 GUI를 구성하여 상기 제스처 데이터를 수집하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 3

제1항에 있어서, 상기 제스처 데이터 수집부(100)는,

라벨링 및 스페이스 바에 의해 제스처의 시작과 끝이 정의된 상기 제스처 데이터를 수집하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 4

제1항에 있어서, 상기 정규화부(200)는,

시간적 정규화 및 공간적 정규화를 실시하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 5

제4항에 있어서, 상기 시간적 정규화는,

보간(Interpolation)과 재표본화(Resampling) 과정을 포함하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 6

제1항에 있어서, 상기 딥 러닝 기반의 3차원 제스처 인식 모델은,

입력 계층(Input Layer), 은닉 계층(Hidden Layer) 및 출력 계층(Output Layer)을 포함하여 구성되는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 7

제6항에 있어서, 상기 입력 계층(Input Layer)은,

100차원의 노드로 구성되고 값이 1인 편향(Bias)을 포함하여 101차원으로 구성되는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 8

제6항에 있어서, 상기 은닉 계층(Hidden Layer)은,

1025개의 노드를 갖는 은닉 계층 1(Hidden Layer 1), 513개의 노드를 갖는 은닉 계층 2(Hidden Layer 2), 257개의 노드를 갖는 은닉 계층 3(Hidden Layer 3), 129개의 노드를 갖는 은닉 계층 4(Hidden Layer 4) 및 65개의 노드를 갖는 은닉 계층 5(Hidden Layer 5)로 구성되어 심층 구조를 형성하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 9

제6항에 있어서, 상기 은닉 계층(Hidden Layer)은,

활성화 함수로 SELU(Scaled Exponential Linear Unit)를 사용하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 10

제6항에 있어서, 상기 출력 계층(Output Layer)은,

21개의 노드로 구성되는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10).

청구항 11

제스처 인식 방법으로서,

(1) 모션 입력 장치를 활용하여 실시간으로 입력된 제스처 데이터를 수집하는 단계;

(2) 상기 단계 (1)에 의해 수집된 제스처 데이터에 대해 정규화를 실시하는 단계;

(3) 상기 단계 (1)에 의해 수집된 제스처 데이터와 상기 단계 (2)에 의해 정규화 된 데이터를 저장하는 단계; 및

(4) 상기 단계 (3)에 의해 저장된 데이터를 이용해 딥 러닝(Deep Learning)에 기초해 학습하여 딥 러닝 기반의 3차원 제스처 인식 모델을 생성하는 단계를 포함하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

청구항 12

제11항에 있어서, 상기 단계 (1)에서는,

상기 모션 입력 장치를 활용하고 게임 엔진을 통해서 GUI를 구성하여 상기제스처 데이터를 수집하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

청구항 13

제11항에 있어서, 상기 단계 (1)에서는,

라벨링 및 스페이스 바에 의해 제스처의 시작과 끝이 정의된 상기 제스처 데이터를 수집하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

청구항 14

제11항에 있어서, 상기 단계 (2)에서는,

시간적 정규화 및 공간적 정규화를 실시하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

청구항 15

제14항에 있어서, 상기 시간적 정규화는,

보간(Interpolation)과 재표본화(Resampling) 과정을 포함하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

청구항 16

제11항에 있어서, 상기 딥 러닝 기반의 3차원 제스처 인식 모델은,

입력 계층(Input Layer), 은닉 계층(Hidden Layer) 및 출력 계층(Output Layer)을 포함하여 구성되는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

청구항 17

제16항에 있어서, 상기 입력 계층(Input Layer)은,

100차원의 노드로 구성되고 값이 1인 편향(Bias)을 포함하여 101차원으로 구성되는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

청구항 18

제16항에 있어서, 상기 은닉 계층(Hidden Layer)은,

1025개의 노드를 갖는 은닉 계층 1(Hidden Layer 1), 513개의 노드를 갖는 은닉 계층 2(Hidden Layer 2), 257개의 노드를 갖는 은닉 계층 3(Hidden Layer 3), 129개의 노드를 갖는 은닉 계층 4(Hidden Layer 4) 및 65개의 노드를 갖는 은닉 계층 5(Hidden Layer 5)로 구성되어 심층 구조를 형성하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

청구항 19

제16항에 있어서, 상기 은닉 계층(Hidden Layer)은,

활성화 함수로 SELU(Scaled Exponential Linear Unit)를 사용하는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

청구항 20

제16항에 있어서, 상기 출력 계층(Output Layer)은,

21개의 노드로 구성되는 것을 특징으로 하는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 제스처 인식 시스템 및 방법에 관한 것으로서, 특히 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템 및 방법에 관한 것이다.

배 경 기 술

[0002] 의사소통을 위해서 인간은 언어를 활용하는 것 외에도 손동작이나 몸동작 같은 제스처를 활용함으로써 쉽게 정보를 전달할 수 있다. 특히, 컴퓨터를 통해서 구축한 가상 환경에서는 제스처를 사용함으로써 사람과 사람, 컴퓨터와 사람 사이에서 직관적으로 정보를 전달하고 명령할 수 있다.

[0003] 제스처를 입력하기 위해서 다양한 입력기기가 필요하며, 크게 2차원 기기와 3차원 기기로 나누어 질 수 있다. 대표적으로 태블릿, 스마트 폰과 같은 모바일 기기의 경우 터치스크린을 사용하여 싱글 또는 멀티 터치 기반으로 2차원 공간에서 제스처가 인식될 수 있다. 또한, 키넥트, 립 모션과 같은 경우에는 3차원 위치 정보를 이용해 제스처의 입력이 가능하다.

[0004] 제스처 인식은 많은 플랫폼에서 활용되고 있으며 입력하는 방식 또한 적용환경에 따라 적절한 형태를 채택할 수 있도록 발전되었다. 특히, 가상의 3차원 공간 안에서는 기존의 입력 형태인 마우스, 키보드의 불편함을 대체하고 효과적인 상호작용 형태가 요구됨에 따라 제스처 인식에 대한 필요성이 대두되었다. 기존 입력 방식은 2차원 환경의 특수성에 최적화되었을 뿐 아니라, 3차원 환경에서도 몰입도와 직관성을 떨어뜨리기 때문에 새로운 입력방법이 필요하게 되었다. 또한, 3차원 환경이 아니더라도 특수한 게임 인터페이스를 가지는 환경에서는 마우스의 조작 방식을 대체하기 위해서 립 모션 등의 기기를 활용해 효과적으로 게임에 몰입할 수 있다. 이러한 특징은 제스처의 직관적인 패턴과 게임의 상황을 자연스럽게 결합한다면 마우스 보다 효과적인 조작 방식이 될 수 있다.

[0005] 제스처 인식은 결정 트리(Decision Tree), 랜덤 포레스트(Random Forest), SVM(Support Vector Machine), 신경망(Neural Network) 계열 알고리즘 등의 기계학습에서 파생된 다양한 알고리즘을 활용할 수 있다. 하지만 제스처 인식은 상대적으로 하드웨어 성능을 크게 요구하며, 연산 속도가 느린 단점이 있다. 반면에 적은 자원을 필요로 하는 알고리즘은 인식률이 떨어져 효과적으로 제스처 인식을 수행할 수 없다.

[0006] 제스처 인식 알고리즘의 성능을 높이기 위해서는 적절한 하드웨어 성능이 요구된다. 게임 어플리케이션과 같이 실시간으로 렌더링 작업을 수행하는 시스템에서는 높은 인식률을 위해 알고리즘의 요구 조건을 만족하기 쉽지 않다. 일반적으로 알고리즘의 속도를 개선하기 위해서는 Open CL, CUDA를 활용해 병렬 처리하는 방식을 선택할 수 있다. 하지만 실시간으로 GPU를 사용하여 렌더링을 처리하는 프로그램에서는 이를 적용하기가 쉽지 않다. 또한, 제스처를 입력하는 형태, 환경 그리고 데이터의 가변성에 따라 알고리즘의 속도와 성능 평가가 달라질 수 있다.

[0007] 최근 신경망을 기반으로 한 딥 러닝은 빅 데이터의 지원, 하드웨어 성능이 발달함에 따라 크게 주목받고 있다. 딥 러닝 알고리즘은 GPU를 활용하여 데이터를 빠르게 연산을 할 수 있으며 효과적으로 데이터를 학습함으로써

다양한 변환에도 강건함을 유지한다. 딥 러닝의 장점은 한 분야에서만 국한되는 것이 아니라 음성 인식, 자연어 처리, 영상 및 이미지 분류에서도 넓게 사용되고 있다. 딥 러닝은 다양한 분야에서 돌파구로 인정되고 있으며 제스처 인식 분야에 적용함으로써 앞서 기술된 문제들을 해결할 수 있을 것이다. 하지만 딥 러닝 알고리즘은 높은 연산량, 소규모 데이터 학습에 대한 과적합 현상과 같은 문제를 갖고 있으며, 이를 해결한다면 효과적으로 제스처 인식을 수행할 수 있을 것이다.

[0008] 한편, 본 발명과 관련된 선행기술로서, 등록특허 제10-1855168(발명의 명칭: 딥러닝 기반의 감성 분류 장치 및 그 방법) 등이 개시된 바 있다.

발명의 내용

해결하려는 과제

[0009] 본 발명은 기존에 제안된 방법들의 상기와 같은 문제점들을 해결하기 위해 제안된 것으로서, 실시간 처리가 필요한 계산 효율 및 성능을 얻기 위해 입력 벡터에 대하여 시간적 및 공간적 정규화를 실시하여 데이터의 손실률을 낮추고, 제스처를 그리는 3차원 공간의 정보를 그대로 활용하여 신뢰도와 정확도를 높일 수 있는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템 및 방법을 제공하는 것을 그 목적으로 한다.

[0010] 또한, 본 발명은, 심층 구조를 적용하고 활성화 함수 SELU(Scaled Exponential Linear Unit)를 사용하여, 빠른 학습 성능과 높은 인식률을 갖고, 시간적 및 공간적 정규화 방법을 통해 딥 러닝 기반의 제스처 인식 모델을 학습하며, 더 나아가 소규모의 데이터에서도 빠르게 학습할 수 있는, 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템 및 방법을 제공하는 것을 또 다른 목적으로 한다.

과제의 해결 수단

[0011] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템은,

[0012] 제스처 인식 시스템으로서,

[0013] 모션 입력 장치를 활용하여 실시간으로 입력된 제스처 데이터를 수집하는 제스처 데이터 수집부;

[0014] 상기 수집된 제스처 데이터에 대해 정규화를 실시하는 정규화부;

[0015] 상기 수집된 제스처 데이터와 상기 정규화부에 의해 정규화 된 데이터를 저장하는 데이터 저장부; 및

[0016] 상기 데이터 저장부에 의해 저장된 데이터를 이용해 딥 러닝(Deep learning)에 기초해 학습하여 딥 러닝 기반의 3차원 제스처 인식 모델을 생성하는 모델 생성부를 포함하는 것을 그 구성상의 특징으로 한다.

[0017] 바람직하게는, 상기 제스처 데이터 수집부는,

[0018] 상기 모션 입력 장치를 활용하고 게임 엔진을 통해서 GUI를 구성하여 상기제스처 데이터를 수집할 수 있다.

[0019] 바람직하게는, 상기 제스처 데이터 수집부는,

[0020] 라벨링 및 스페이스 바에 의해 제스처의 시작과 끝이 정의된 상기 제스처 데이터를 수집할 수 있다.

[0021] 바람직하게는, 상기 정규화부는,

[0022] 시간적 정규화 및 공간적 정규화를 실시할 수 있다.

[0023] 더욱 바람직하게는, 상기 시간적 정규화는,

[0024] 보간(Interpolation)과 재표본화(Resampling) 과정을 포함할 수 있다.

[0025] 바람직하게는, 상기 딥 러닝 기반의 3차원 제스처 인식 모델은,

[0026] 입력 계층(Input Layer), 은닉 계층(Hidden Layer) 및 출력 계층(Output Layer)을 포함하여 구성될 수 있다.

[0027] 더욱 바람직하게는, 상기 입력 계층(Input Layer)은,

[0028] 100차원의 노드로 구성되고 값이 1인 편향(Bias)을 포함하여 101차원으로 구성될 수 있다.

[0029] 더욱 바람직하게는, 상기 은닉 계층(Hidden Layer)은,

[0030] 1025개의 노드를 갖는 은닉 계층 1(Hidden Layer 1), 513개의 노드를 갖는 은닉 계층 2(Hidden Layer 2), 257개의 노드를 갖는 은닉 계층 3(Hidden Layer 3), 129개의 노드를 갖는 은닉 계층 4(Hidden Layer 4) 및 65개의 노드를 갖는 은닉 계층 5(Hidden Layer 5)로 구성되어 심층 구조를 형성할 수 있다.

[0031] 더욱 바람직하게는, 상기 은닉 계층(Hidden Layer)은,

[0032] 활성화 함수로 SELU(Scaled Exponential Linear Unit)를 사용할 수 있다.

[0033] 더욱 바람직하게는, 상기 출력 계층(Output Layer)은,

[0034] 21개의 노드로 구성될 수 있다.

[0035] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법은,

[0036] 제스처 인식 방법으로서,

[0037] (1) 립 모션을 활용하여 복수의 피실험자로부터 입력된 제스처 데이터를 수집하는 단계;

[0038] (2) 상기 단계 (1)에 의해 수집된 제스처 데이터에 대해 정규화를 실시하는 단계;

[0039] (3) 입력된 제스처 데이터를 3차원 손 위치 데이터와 상기 단계 (2)에 의해 정규화 된 데이터를 저장하는 단계; 및

[0040] (4) 상기 단계 (3)에 의해 저장된 데이터를 이용해 딥 러닝(Deep Learning)에 기초해 학습하여 딥 러닝 기반의 3차원 제스처 인식 모델을 생성하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

[0041] 바람직하게는, 상기 단계 (1)에서는,

[0042] 상기 모션 입력 장치를 활용하고 게임 엔진을 통해서 GUI를 구성하여 상기 제스처 데이터를 수집할 수 있다.

[0043] 바람직하게는, 상기 단계 (1)에서는,

[0044] 라벨링 및 스페이스 바에 의해 제스처의 시작과 끝이 정의된 상기 제스처 데이터를 수집할 수 있다.

[0045] 바람직하게는, 상기 단계 (2)에서는,

[0046] 시간적 정규화 및 공간적 정규화를 실시할 수 있다.

[0047] 더욱 바람직하게는, 상기 시간적 정규화는,

[0048] 보간(Interpolation)과 재표본화(Resampling) 과정을 포함할 수 있다.

[0049] 바람직하게는, 상기 딥 러닝 기반의 3차원 제스처 인식 모델은,

[0050] 입력 계층(Input Layer), 은닉 계층(Hidden Layer) 및 출력 계층(Output Layer)을 포함하여 구성될 수 있다.

[0051] 더욱 바람직하게는, 상기 입력 계층(Input Layer)은,

[0052] 100차원의 노드로 구성되고 값이 1인 편향(Bias)을 포함하여 101차원으로 구성될 수 있다.

[0053] 더욱 바람직하게는, 상기 은닉 계층(Hidden Layer)은,

[0054] 1025개의 노드를 갖는 은닉 계층 1(Hidden Layer 1), 513개의 노드를 갖는 은닉 계층 2(Hidden Layer 2), 257개의 노드를 갖는 은닉 계층 3(Hidden Layer 3), 129개의 노드를 갖는 은닉 계층 4(Hidden Layer 4) 및 65개의 노드를 갖는 은닉 계층 5(Hidden Layer 5)로 구성되어 심층 구조를 형성할 수 있다.

[0055] 더욱 바람직하게는, 상기 은닉 계층(Hidden Layer)은,

[0056] 활성화 함수로 SELU(Scaled Exponential Linear Unit)를 사용할 수 있다.

[0057] 더욱 바람직하게는, 상기 출력 계층(Output Layer)은,

[0058] 21개의 노드로 구성될 수 있다.

발명의 효과

[0059] 본 발명에서 제안하고 있는 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템 및 방법에 따르면, 실시간 처리가 필요한 계산 효율 및 성능을 얻기 위해 입력 벡터에 대하여 시간적 및 공간적 정규화를 실시하여 데이터의 손실률을 낮추고, 제스처를 그리는 3차원 공간의 정보를 그대로 활용하여 신뢰도와 정확도를 높일 수 있다.

[0060] 또한, 본 발명에 따르면, 심층 구조를 적용하고 활성화 함수 SELU(Scaled Exponential Linear Unit)를 사용하여, 빠른 학습 성능과 높은 인식률을 갖고, 시간적 및 공간적 정규화 방법을 통해 딥 러닝 기반의 제스처 인식 모델을 학습하며, 더 나아가 소규모의 데이터에서도 빠르게 학습할 수 있다.

도면의 간단한 설명

[0061] 도 1은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 구성을 블록도로 도시한 도면.

도 2는 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 제스처 데이터 수집부에서 모션 입력 장치를 활용하여 제스처 데이터를 수집하고 있는 모습을 도시한 도면.

도 3은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 정규화부에서 이루어지는 정규화 과정을 나타낸 도면.

도 4는 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식

시스템의 정규화부에서 제스처 데이터를 입력받는 과정을 설명하기 위해 도시한 도면.

도 5는 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 정규화부에서 이루어지는 보간(Interpolation) 과정을 설명하기 위해 도시한 도면.

도 6은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 정규화부에서 이루어지는 재표본화(Resampling) 과정을 설명하기 위해 도시한 도면.

도 7은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 정규화부에서 이루어지는 바운딩 박스(Bounding Box)를 생성하는 과정을 설명하기 위해 도시한 도면.

도 8은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 정규화부에서 바운딩 박스(Bounding Box)를 생성한 후, 정규화를 실시하는 과정을 설명하기 위해 도시한 도면.

도 9는 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 데이터 저장부에서 제스처 데이터를 저장하는 과정을 설명하기 위해 도시한 도면.

도 10은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 데이터 저장부에서 저장 과정을 거친 제스처 데이터의 (a) 손 위치 데이터가 저장된 모습과, (b) 정규화 된 데이터가 저장된 모습을 도시한 도면.

도 11은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 모델 생성부에서 생성된 딥 러닝 기반의 3차원 제스처 인식 모델의 구조를 도시한 도면.

도 12는 본 발명의 일실시예에 따른 딥 러닝 기반의 3차원 제스처 인식 모델의 은닉 계층에서 활성화 함수로 활용하는 SELU(Scaled Exponential Linear Unit)를 설명하기 위해 도시한 도면.

도 13은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법의 흐름을 도시한 도면.

발명을 실시하기 위한 구체적인 내용

[0062] 이하에서는 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 동일 또는 유사한 부호를 사용한다.

[0063] 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.

[0064] 도 1은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 구성을 블록도로 도시한 도면이다. 도 1에 도시된 바와 같이, 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)은, 제스처 데이터 수집부(100), 정규화부(200), 데이터 저장부(300) 및 모델 생성부(400)를 포함하여 구성될 수 있다. 이하에서는, 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 각각의 구성에 대하여 상세히 설명하기로 한다.

[0065] 제스처 데이터 수집부(100)는, 모션 입력 장치를 활용하여 실시간으로 입력된 제스처 데이터를 수집할 수 있다. 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 제스처 데이터 수집부(100)는, 모션 입력 장치를 활용하고 게임 엔진을 통해서 GUI를 구성하여 제스처

데이터를 수집할 수 있다. 이때, 모션 입력 장치는 3차원 제스처를 인식할 수 있는 기기로서, 키넥트, 립 모션이 될 수 있다. 다만, 상기의 키넥트, 립 모션으로 모션 입력 장치를 한정하는 것은 아니다. 또한, 게임엔진으로서 Unity 5 게임엔진을 사용할 수 있으나, 상기의 예로 한정하는 것은 아니다.

[0066] 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 제스처 데이터 수집부(100)는, 라벨링 및 스페이스 바에 의해 제스처의 시작과 끝이 정의된 제스처 데이터를 수집할 수 있다. 예를 들면, 제스처를 시작할 때 스페이스바를 눌러 사용자의 손 움직임으로부터 제스처 데이터를 입력 받을 수 있고 제스처가 끝나면 스페이스바를 눌러 제스처 데이터를 입력 받는 것을 끝마칠 수 있다.

[0067] 다만, 많은 제스처 데이터를 사람이 직접 라벨링 작업을 수행하기 어려우므로, 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)에서는 별도로 라벨링 작업을 수행하는 프로그램을 구축하여 라벨링 작업을 수행할 수 있다.

[0068] 정규화부(200)는, 제스처 데이터 수집부(100)에 의해 수집된 제스처 데이터에 대해 정규화를 실시할 수 있다. 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)에서 정규화부(200)는, 사용자의 제스처 속도에 강건한 제스처 데이터를 얻기 위한 시간적 정규화와, 제스처의 크기나 이동에 불변하는 제스처 데이터를 얻기 위한 공간적 정규를 사용하여 정규화를 실시할 수 있다. 이때, 시간적 정규화에는 보간(Interpolation)과 재표본화(Resampling) 과정이, 공간적 정규화에는 바운딩 박스(Bounding Box)를 생성하는 과정이 포함될 수 있다.

[0069] 도 3은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 정규화부에서 이루어지는 정규화 과정을 나타낸 도면이다. 도 3에 도시된 바와 같이, 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 정규화부(200)는, 제스처 데이터 입력, 보간(Interpolation), 재표본화(Resampling) 및 바운딩 박스(Bounding Box) 생성 과정을 통해 시간적 및 공간적 정규화를 실시할 수 있다. 이하에서는 각각의 과정들에 대해 상세히 설명하기로 한다.

[0070] 도 4는 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 정규화부에서 제스처 데이터를 입력받는 과정을 설명하기 위해 도시한 도면이다. 도 4에 도시된 바와 같이, 정규화 과정은 제스처 데이터를 벡터 형태로 입력 받아 실시될 수 있다. 이때, 제스처 데이터는 초당 프레임 수마다 벡터로 입력될 수 있다. 다만, 제스처 데이터를 벡터 형태로 입력받는 과정에서 입력 벡터가 소실될 수 있다. 입력 벡터의 소실을 막기 위하여 입력받은 벡터 간의 보간(Interpolation)을 수행할 수 있다.

[0071] 도 5는 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 정규화부에서 이루어지는 보간(Interpolation) 과정을 설명하기 위해 도시한 도면이다. 도 5에 도시된 바와 같이, 입력 데이터는 3차원 벡터이며, 데이터를 입력받을 때마다 벡터의 소실을 막기 위하여 아래의 수학식 1에 따라 두 벡터 사이의 선형 보간(Linear Interpolation)을 수행할 수 있다.

수학식 1

[0072] $$x = [0,1], v(x) = (1-x)v(t-1) + v(t)$$

[0073] 보간 구간은 [0,1]이며, 이전 입력 받은 벡터 v(t-1)와 현재 입력 받은 벡터 v(t)에 대해 보간을 할 수 있다. 이때, 보간 간격과 보간 벡터의 생성 수는 컴퓨터의 작업 환경을 고려하여, 상황에 따라 적합한 수치를 선정할

수 있다.

[0074] 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 정규화부(200)는 딥 러닝 기반의 3차원 제스처 인식 모델의 적합한 차원의 입력 데이터를 생성하기 위해 입력받은 벡터를 소거하거나 보충하는 재표본화(Resampling) 과정을 수행할 수 있다. 도 6은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 정규화부에서 이루어지는 재표본화(Resampling) 과정을 설명하기 위해 도시한 도면이다. 도 6에 도시된 바와 같이, 입력 데이터 차원 보다 벡터가 적을 경우에는 보간하는 과정과 입력 데이터 차원 보다 벡터가 많을 경우 소거하는 과정을 수행할 수 있다.

[0075] 이때, 벡터를 소거하는 과정은 현재까지 입력받은 벡터의 수와 데이터 처리 전의 제스처 인식 모델이 요구하는 벡터의 수를 나누어서 간격을 구한 후, 도출된 간격에 따라 입력받은 벡터 중 보존해야 할 대상을 선정하고 선정되지 못한 벡터를 소거하는 것이고, 벡터를 보충하는 과정은 딥 러닝 기반의 3차원 제스처 인식 모델이 요구하는 입력 차원을 획득하지 못하고 입력 데이터가 누락된 경우 입력 벡터를 보충하는 것이다.

[0076] 정규화부(200)는, 소거 및 보충하는 과정을 거친 입력 벡터에 대해 정규화를 실시하기 위해 바운딩 박스(Bounding Box)를 생성할 수 있다. 도 7은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템의 정규화부에서 이루어지는 바운딩 박스(Bounding Box)를 생성하는 과정을 설명하기 위해 도시한 도면이다. 도 7에 도시된 바와 같이, 전체 입력 벡터로부터 가장 큰 원소를 추출한 벡터 v_{max}와 전체 입력 벡터로부터 가장 작은 원소를 추출한 벡터 v_{min}을 도출하여 바운딩 박스(Bounding Box)를 생성할 수 있다. 도출된 벡터를 통해서 각 입력 벡터에 대해서 정규화를 할 수 있다.

[0077] 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 정규화부(200)는 바운딩 박스(Bounding Box) 생성 후, 각 입력 벡터에 대한 정규화 과정을 실시할 수 있다. 도 8은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 정규화부(200)에서 바운딩 박스(Bounding Box)를 생성한 후, 정규화를 실시하는 과정을 설명하기 위해 도시한 도면이다. 도 8에 도시된 바와 같이, 생성된 바운딩 박스(Bounding Box)를 활용하여 최대 노름(Max norm)을 통해 입력 벡터에 대해 정규화를 할 수 있다.

[0078] 정규화 과정에서는 모든 3차원 벡터가 [0,1]의 범위를 가지기 위해서, 축의 값이 가장 작은 벡터와 큰 벡터를 도출하여 적합한 범위를 가지는 과정을 수행할 수 있다. 바운딩 박스(Bounding Box)를 통해서 정규화 과정을 정리한 식은 수학식 2와 같다.

수학식 2

[0079]
$$v^i = (v_x^i, v_y^i, v_z^i) \in R^3,\ i = 0, \cdots, n$$
$$v_{\min} = (\min_i v_x^i,\ \min_i v_y^i,\ \min_i v_z^i)$$
$$v_{\max} = (\max_i v_x^i,\ \max_i v_y^i,\ \max_i v_z^i)$$
$$v^k = \frac{v_i - v_{\min}}{\|v_{\max} - v_{\min}\|_\infty}$$

[0080] 입력 벡터를 v^i, 전체 입력 벡터로부터 가장 작은 원소를 추출한 벡터를 v_{min}, 전체 입력 벡터로부터 가장 큰 원소를 추출한 벡터를 v_{max}라 할 때, v_{min}과 v^i의 차를 구한 후, 두 벡터 v_{max}, v_{min}의 차에 대한 최대 노름(Max

norm)을 통해서 원소 중 가장 큰 값이 도출된다. 도출된 값을 사용하여 v_{min}과 v^i의 차를 나누는 과정을 수행하고 결과적으로 정규화 된 벡터 v^k를 도출할 수 있다. 정규화 후, 데이터 생성은 3차원 위치 벡터 x, y, z축의 값과 오른손, 왼손을 구분하는 파라미터 1개를 통해서 총 100차원의 벡터를 생성할 수 있다.

[0081] 데이터 저장부(300)는, 제스처 데이터 수집부(100)에 의해 수집된 제스처 데이터와 정규화부(200)에 의해 정규화 된 데이터를 저장할 수 있다. 도 9는 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 데이터 저장부(300)에서 제스처 데이터를 저장하는 과정을 설명하기 위해 도시한 도면이다. 도 9에 도시된 바와 같이, 데이터 저장부(300)는 제스처 데이터를 제스처에 대한 라벨링 번호를 통해 구분하고, 제스처 데이터 수집부(100)에 의해 수집된 제스처 데이터를 손 위치 데이터로, 정규화부(200)에 의해 정규화 된 데이터로 각각 다른 폴더에 새로운 포맷 형태로 저장할 수 있다. 이때, 폴더의 이름은 raw와 preprocess가, 포맷의 형태는 vec와 d3g가 될 수 있으나, 상기의 예로 한정하는 것은 아니다.

[0082] 도 10은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 데이터 저장부(300)에서 저장 과정을 거친 제스처 데이터의 (a) 손 위치 데이터가 저장된 모습과, (b) 정규화 된 데이터가 저장된 모습을 도시한 도면이다. 도 10에 도시된 바와 같이, 손 위치 데이터는 3차원 벡터 x, y, z 값이 원형으로 저장될 수 있고, 정규화 된 데이터는 3차원 벡터에 대해서 정규화 과정을 거치기 때문에 [0,1]사이의 데이터로 저장될 수 있다.

[0083] 모델 생성부(400)은, 데이터 저장부(300)에 의해 저장된 데이터를 이용해 딥 러닝(Deep Learning)에 기초해 학습하여 딥 러닝 기반의 3차원 제스처 인식 모델을 생성할 수 있다. 도 11은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)의 모델 생성부(400)에서 생성된 딥 러닝 기반의 3차원 제스처 인식 모델의 구조를 도시한 도면이다. 도 11에 도시된 바와 같이, 딥 러닝 기반의 3차원 제스처 인식 모델은 입력 계층(Input Layer), 은닉 계층(Hidden Layer) 및 출력 계층(Output Layer)을 포함하여 구성될 수 있다. 첫 번째 입력 계층은 100차원의 노드로 구성되고 값이 1인 편향(Bias)을 포함하여 101차원으로 구성될 수 있다. 두 번째 은닉 계층은 1025개의 노드를 갖는 은닉 계층 1(Hidden Layer 1), 513개의 노드를 갖는 은닉 계층 2(Hidden Layer 2), 257개의 노드를 갖는 은닉 계층 3(Hidden Layer 3), 129개의 노드를 갖는 은닉 계층 4(Hidden Layer 4) 및 65개의 노드를 갖는 은닉 계층 5(Hidden Layer 5)로 구성되어 심층 구조를 형성할 수 있다. 마지막으로 출력 계층은 21개의 노드로 구성될 수 있다.

[0084] 도 12는 본 발명의 일실시예에 따른 딥 러닝 기반의 3차원 제스처 인식 모델의 은닉 계층에서 활성화 함수로 활용하는 SELU(Scaled Exponential Linear Unit)를 설명하기 위해 도시한 도면이다. 도 12에 도시된 바와 같이, 본 발명의 일실시예에 따른 딥 러닝 기반의 3차원 제스처 인식 모델의 은닉 계층은 활성화 함수로 SELU(Scaled Exponential Linear Unit)를 사용할 수 있다. 이때, SELU의 식은 수학식 3과 같다.

수학식 3

[0085]
$$f(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & otherwise \end{cases}$$

[0086] SELU는 입력 값이 양수일 때는 λ를 곱해서 출력되고, 음수일 때는 ($\alpha e^x - \alpha$)의 값에 λ를 곱해서 출력할 수 있다. 활성화 함수 SELU를 통해서 기존 ReLU 함수가 학습 시 누적된 가중치가 폭발하는 현상을 방지할 수 있으며, 과적합을 부분적으로 방지하여 학습 성능을 높일 수 있다.

[0087] 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)은, 3차원 공간상의 정보를 그대로 활용하여 팔 부위의 재활치료를 실시할 수 있다. 같은 종류의 제스처이

더라도 재활치료 환자를 위한 동작을 필요로 한다는 것을 감안하고 팔과 손이 불편한 환자와 유사한 데이터를 수집하여, 데이터의 편향성이 없고 신뢰도가 높은 데이터 집합을 구축함으로써, 실제 재활 환자들로 하여금 지루하지 않은 재활치료가 가능하도록 하는 환경을 만들 수 있다.

[0088] 도 13은 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법의 흐름을 도시한 도면이다. 도 13에 도시된 바와 같이, 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 방법은, 실시간 3차원 제스처 인식 방법으로서, 모션 입력 장치를 활용하여 실시간으로 입력된 제스처 데이터를 수집하는 단계(S100), 상기 단계 S100에 의해 수집된 제스처 데이터에 대해 정규화를 실시하는 단계(S200), 상기 단계 S100에 의해 수집된 제스처 데이터와 상기 단계 S200에 의해 정규화 된 데이터를 저장하는 단계(S300), 및 상기 단계 S300에 의해 저장된 데이터를 이용해 딥 러닝(Deep Learning)에 기초해 학습하여 딥 러닝 기반의 3차원 제스처 인식 모델을 생성하는 단계(S400)를 포함하여 구성될 수 있다.

[0089] 각각의 단계들과 관련된 상세한 내용들은, 앞서 본 발명의 일실시예에 따른 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템(10)과 관련하여 충분히 설명되었으므로, 상세한 설명은 생략하기로 한다.

[0090] 전술한 바와 같이, 본 발명에서 제안하고 있는 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템 및 방법에 따르면, 실시간 처리가 필요한 계산 효율 및 성능을 얻기 위해 입력 벡터에 대하여 시간적 및 공간적 정규화를 실시하여 데이터의 손실률을 낮추고, 제스처를 그리는 3차원 공간의 정보를 그대로 활용하여 신뢰도와 정확도를 높일 수 있다. 또한, 본 발명에 따르면, 심층 구조를 적용하고 활성화 함수 SELU(Scaled Exponential Linear Unit)를 사용하여, 빠른 학습 성능과 높은 인식률을 갖고, 시간적 및 공간적 정규화 방법을 통해 딥 러닝 기반의 제스처 인식 모델을 학습하며, 더 나아가 소규모의 데이터에서도 빠르게 학습할 수 있다.

[0091] 이상 설명한 본 발명은 본 발명이 속한 기술분야에서 통상의 지식을 가진 자에 의하여 다양한 변형이나 응용이 가능하며, 본 발명에 따른 기술적 사상의 범위는 아래의 특허청구범위에 의하여 정해져야 할 것이다.

부호의 설명

[0092] 10: 시간적 및 공간적 정규화를 통한 딥 러닝 기반의 실시간 3차원 제스처 인식 시스템

100: 제스처 데이터 수집부

200: 정규화부

300: 데이터 저장부

400: 모델 생성부

S100: 모션 입력 장치를 활용하여 실시간으로 입력된 제스처 데이터를 수집하는 단계

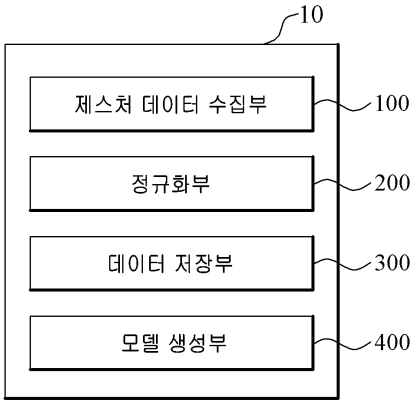
S200: 단계 S100에 의해 수집된 제스처 데이터에 대해 정규화를 실시하는 단계

S300: 단계 S100에 의해 수집된 제스처 데이터와 단계 S200에 의해 정규화 된 데이터를 저장하는 단계

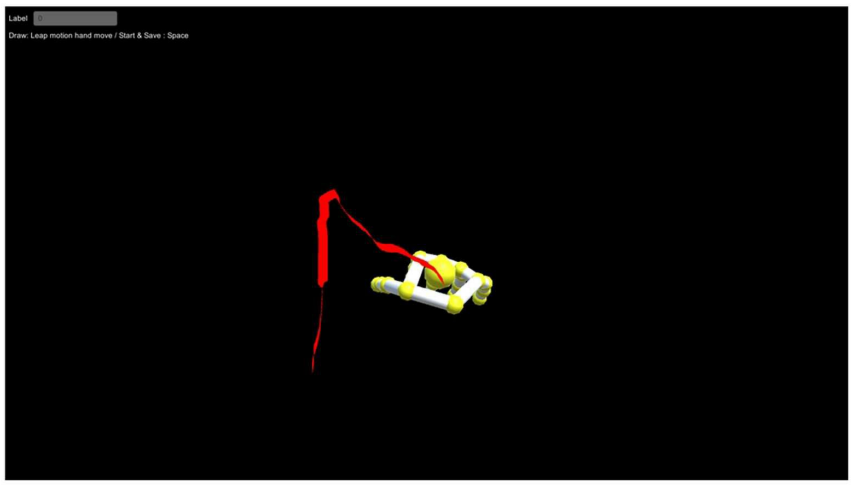
S400: 단계 S300에 의해 저장된 데이터를 이용해 딥 러닝(Deep Learning)에 기초해 학습하여 딥 러닝 기반의 3차원 제스처 인식 모델을 생성하는 단계

도면

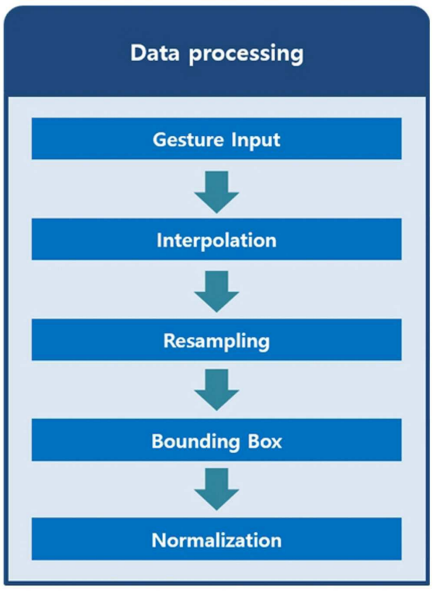
도면1



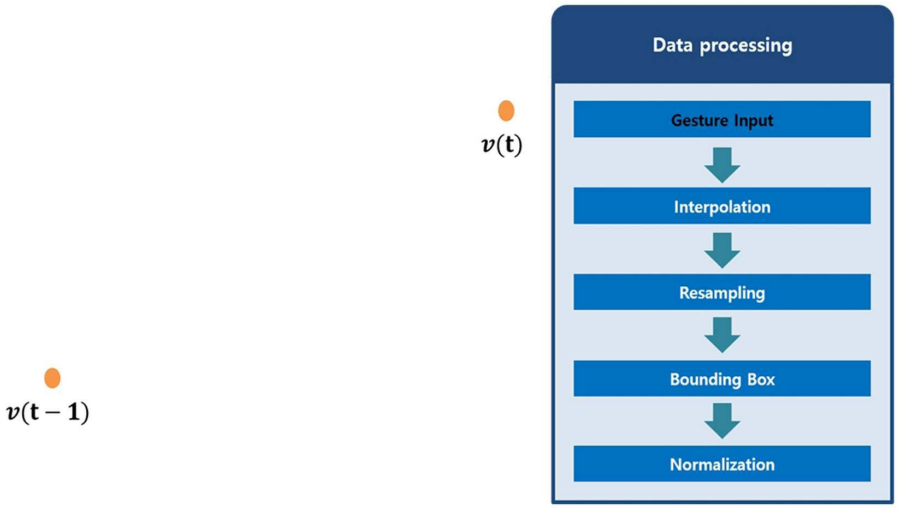
도면2



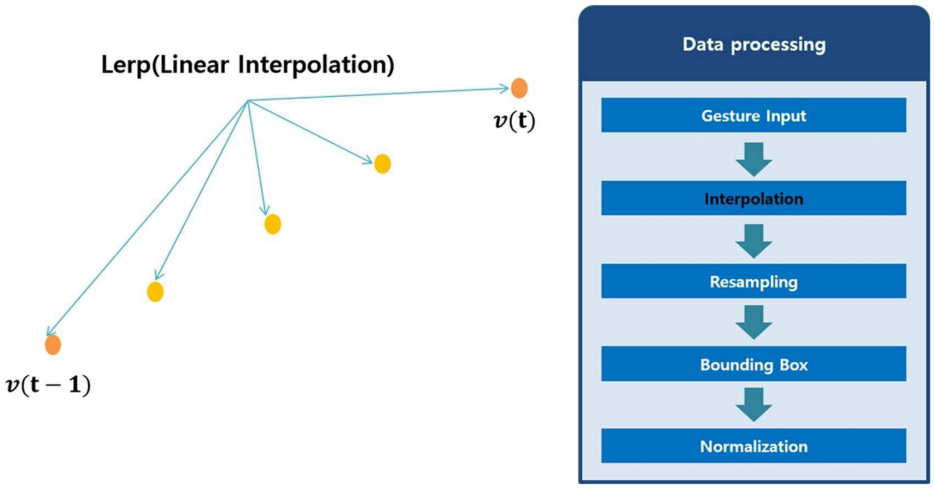
도면3



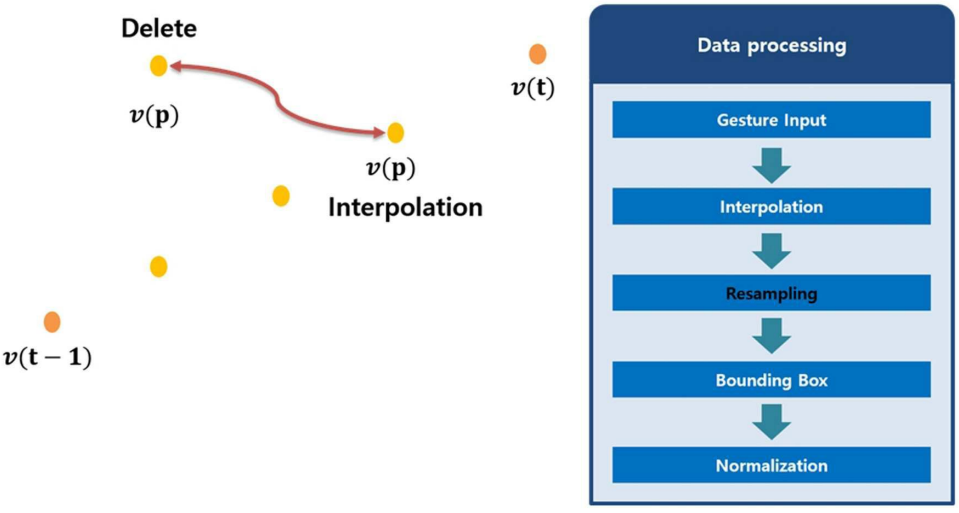
도면4


Data processing
Gesture Input
Interpolation
Resampling
Bounding Box
Normalization
v(t)
v(t − 1)

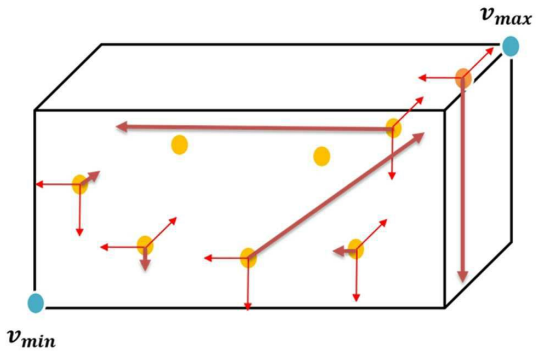
도면5

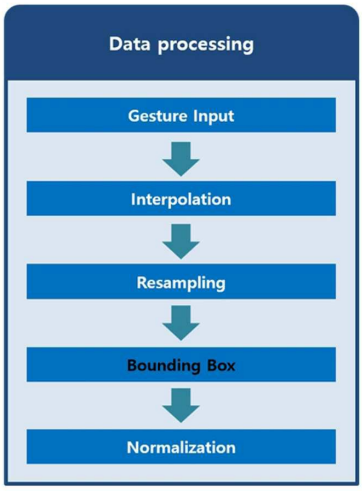

Lerp(Linear Interpolation)
v(t)
v(t − 1)
Data processing
Gesture Input
Interpolation
Resampling
Bounding Box
Normalization

도면6

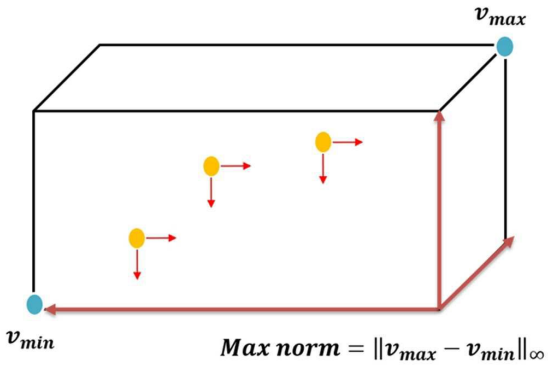

Delete
v(p)
v(p)
Interpolation
v(t)
v(t − 1)
Data processing
Gesture Input
Interpolation
Resampling
Bounding Box
Normalization

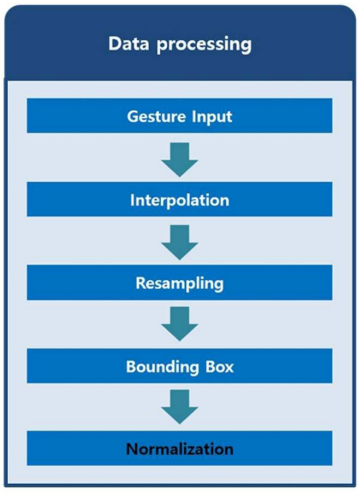
도면7


v_{max}
v_{min}


Data processing
Gesture Input
Interpolation
Resampling
Bounding Box
Normalization

도면8


v_{max}
v_{min}
$Max\ norm = \|v_{max} - v_{min}\|_\infty$


Data processing
Gesture Input
Interpolation
Resampling
Bounding Box
Normalization

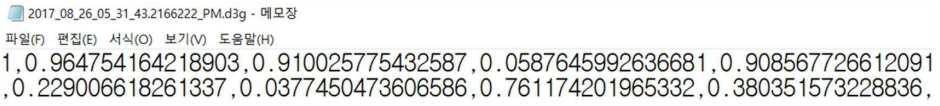
도면9


0
1
2
preprocess
raw
2017_08_26_05_31_29.1245952_PM.vec
2017_08_26_05_31_30.7716699_PM.vec
2017_08_26_05_31_32.3165029_PM.vec
2017_08_26_05_31_33.7517077_PM.vec
2017_08_26_05_31_35.1309713_PM.vec
2017_08_26_05_31_36.4942166_PM.vec
2017_08_26_05_31_37.8505289_PM.vec
2017_08_26_05_31_39.1356580_PM.vec
2017_08_26_05_31_40.4873809_PM.vec
2017_08_26_05_31_41.8816044_PM.vec
2017_08_26_05_31_43.2166222_PM.vec
2017_08_26_05_31_29.1245952_PM.d3g
2017_08_26_05_31_30.7716699_PM.d3g
2017_08_26_05_31_32.3165029_PM.d3g
2017_08_26_05_31_33.7517077_PM.d3g
2017_08_26_05_31_35.1309713_PM.d3g
2017_08_26_05_31_36.4942166_PM.d3g
2017_08_26_05_31_37.8505289_PM.d3g
2017_08_26_05_31_39.1356580_PM.d3g
2017_08_26_05_31_40.4873809_PM.d3g
2017_08_26_05_31_41.8816044_PM.d3g
2017_08_26_05_31_43.2166222_PM.d3g

도면10

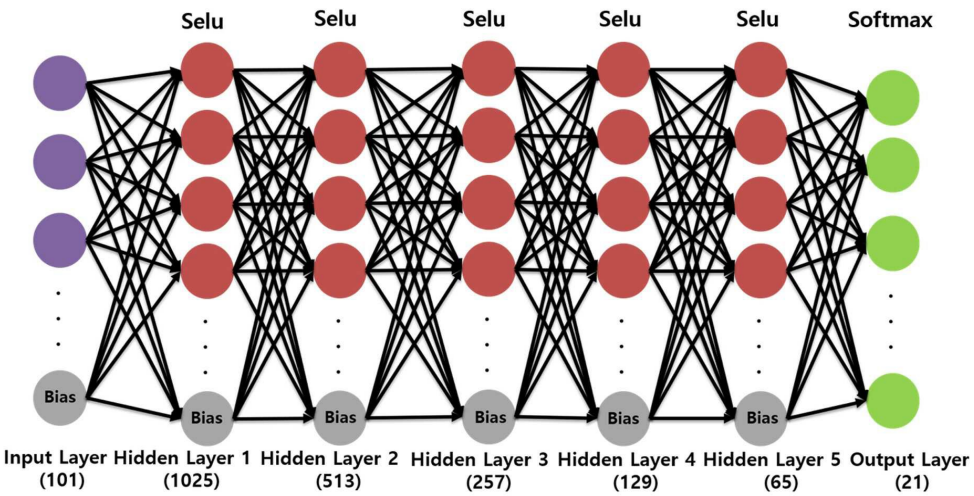

2017_08_26_05_31_43.2166222_PM.vec - 메모장
파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)
-0.1695887,0.06904849,-9.543911
-0.1695007,0.06996522,-9.544142
-0.1694126,0.07088195,-9.544374
-0.1693246,0.07179868,-9.544604
-0.1692366,0.07271541,-9.544836
-0.1691485,0.07363214,-9.545067
-0.1690605,0.07454886,-9.545298
-0.1689724,0.07546559,-9.545529
-0.1688844,0.07638232,-9.54576
-0.1687964,0.07729905,-9.545992
-0.1687083,0.07821578,-9.546223
-0.1683684,0.07909898,-9.546398
-0.1680284,0.07998219,-9.546575
-0.1676885,0.0808654,-9.54675

(a)

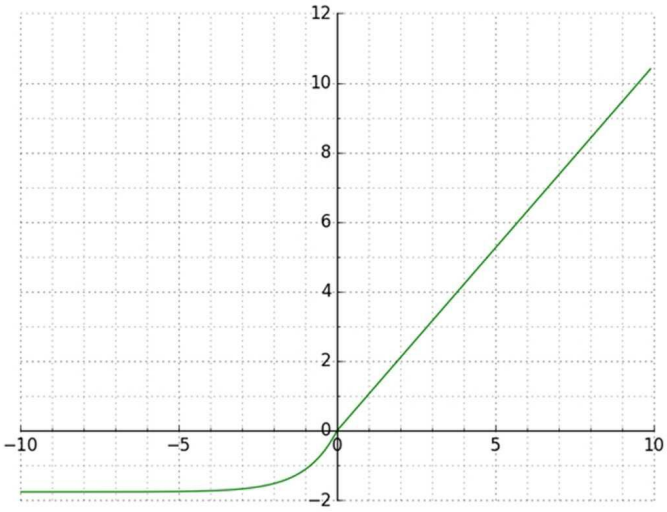

2017_08_26_05_31_43.2166222_PM.d3g - 메모장
파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)
1,0.964754164218903,0.910025775432587,0.0587645992636681,0.908567726612091
,0.229006618261337,0.0377450473606586,0.761174201965332,0.380351573228836,

(b)

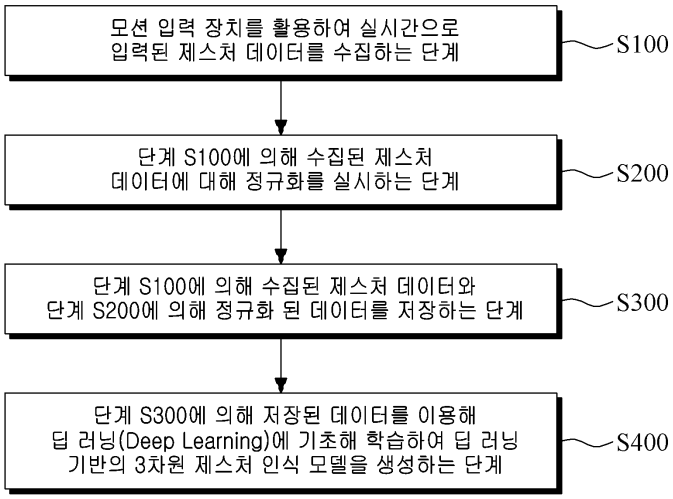
도면11


Selu
Selu
Selu
Selu
Selu
Softmax
Bias
Bias
Bias
Bias
Bias
Bias
Input Layer (101)
Hidden Layer 1 (1025)
Hidden Layer 2 (513)
Hidden Layer 3 (257)
Hidden Layer 4 (129)
Hidden Layer 5 (65)
Output Layer (21)

도면12


12
10
8
6
4
2
0
−2
−10
−5
0
5
10

도면13


모션 입력 장치를 활용하여 실시간으로
입력된 제스처 데이터를 수집하는 단계
S100
단계 S100에 의해 수집된 제스처
데이터에 대해 정규화를 실시하는 단계
S200
단계 S100에 의해 수집된 제스처 데이터와
단계 S200에 의해 정규화 된 데이터를 저장하는 단계
S300
단계 S300에 의해 저장된 데이터를 이용해
딥 러닝(Deep Learning)에 기초해 학습하여 딥 러닝
기반의 3차원 제스처 인식 모델을 생성하는 단계
S400