



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년07월04일
(11) 등록번호 10-2417150
(24) 등록일자 2022년06월30일

(51) 국제특허분류(Int. Cl.)
G06F 3/01 (2006.01) G06N 20/00 (2019.01)
G06V 10/40 (2022.01)
(52) CPC특허분류
G06F 3/011 (2022.02)
G06N 20/00 (2021.08)
(21) 출원번호 10-2019-0148676
(22) 출원일자 2019년11월19일
심사청구일자 2019년11월19일
(65) 공개번호 10-2021-0060986
(43) 공개일자 2021년05월27일
(56) 선행기술조사문헌
KR1020140006665 A
KR1020140051344 A
KR1020150116641 A
KR1020200048032 A

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
이종석
인천광역시 남동구 소래역로 93, 906-104
문성은
서울특별시 구로구 경인로20가길 68, 101-1716
(74) 대리인
김권석

전체 청구항 수 : 총 3 항

심사관 : 임지환

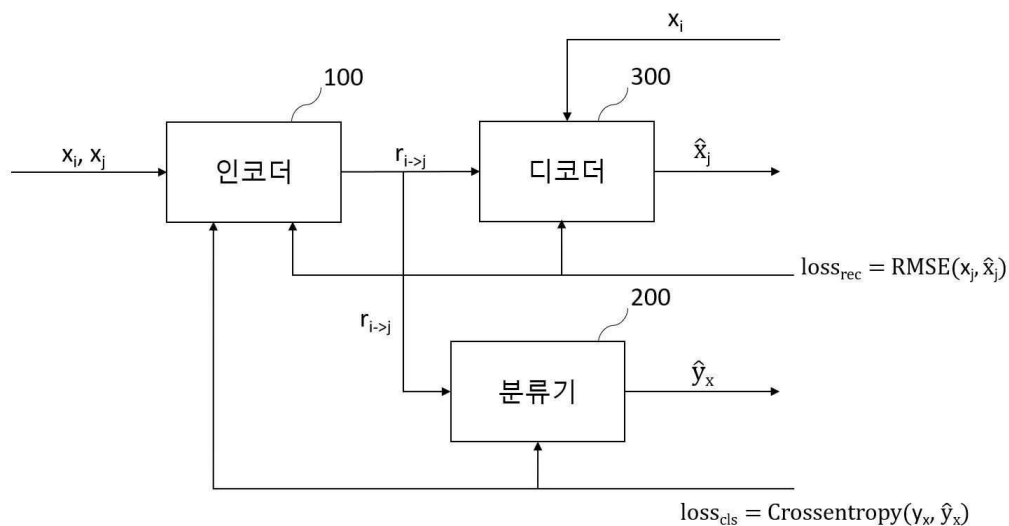
(54) 발명의 명칭 사용자 동작 인식 방법 및 이를 위한 시스템

(57) 요약

본 발명의 일 실시예에 따른 사용자 동작 인식 시스템은, 복수 개의 신호를 입력받는 신호 입력부; 복원 손실 정보, 분류 손실 정보, 및 상기 복수 개의 신호를 수신하여 상기 복수 개의 신호 중 제 1 신호와 제 2 신호와의 관계를 나타내는 특성 정보를 추출하는 특징 추출부; 및 상기 복원 손실 정보를 디코더로부터 수신하는 복원 손실

(뒷면에 계속)

대표도 - 도5



입력부를 포함하는 인코더; 상기 특성 정보를 기초로 하여 분류 인덱스를 예측하는 신호 분류부; 및 상기 예측된 분류 인덱스와 오리지널 분류 인덱스의 분류 손실을 산출하는 분류 손실 산출부를 포함하는 분류기; 및 상기 제 1 신호 및 상기 인코더로부터 상기 특성 정보를 수신하는 신호 수신부; 상기 제 1 신호 및 상기 특성 정보를 이용하여 상기 제 2 신호를 복원하는 신호 복원부; 및 상기 복원된 제 2 신호와 오리지널 제 2 신호를 비교하여 복원 손실을 산출하는 복원 손실 산출부를 포함하는 디코더를 포함하고, 상기 분류 손실 정보는 상기 인코더 및 상기 분류기에 입력되어 상기 인코더 및 상기 분류기를 학습시키고, 상기 복원 손실 정보는 상기 인코더 및 상기 디코더에 입력되어 상기 인코더 및 상기 디코더를 학습시킬 수 있다.

이 발명을 지원한 국가연구개발사업

과제고유번호	2019110679
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원(한국연구재단부설)
연구사업명	정보통신방송연구개발사업
연구과제명	[이지마로][주관/경북대학교] 사용자의 의도와 맥락을 이해하는 지능형 인터랙션 기
술 연구개발(4/5)	
기 여 율	1/1
과제수행기관명	경북대학교
연구기간	2019.05.01 ~ 2020.02.29

명세서

청구범위

청구항 1

삭제

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

복수 개의 신호를 입력받는 신호 입력부; 복원 손실 정보, 분류 손실 정보, 및 상기 복수 개의 신호를 수신하여 상기 복수 개의 신호 중 제 1 신호와 제 2 신호와의 관계를 나타내는 특성 정보를 추출하는 특징 추출부; 및 상기 복원 손실 정보를 디코더로부터 수신하는 복원 손실 입력부를 포함하는 인코더;

상기 특성 정보를 기초로 하여 분류 인덱스를 예측하는 신호 분류부; 및 상기 예측된 분류 인덱스와 오리지널 분류 인덱스의 분류 손실을 산출하는 분류 손실 산출부를 포함하는 분류기; 및

상기 제 1 신호 및 인코더로부터 상기 특성 정보를 수신하는 신호 수신부; 상기 제 1 신호 및 상기 특성 정보를 이용하여 상기 제 2 신호를 복원하는 신호 복원부; 및 상기 복원된 제 2 신호와 오리지널 제 2 신호를 비교하여 복원 손실을 산출하는 복원 손실 산출부를 포함하는 디코더를 포함하고,

상기 분류 손실 정보는 상기 인코더 및 상기 분류기에 입력되어 상기 인코더 및 상기 디코더를 학습시키고,

상기 복원 손실 정보는 상기 인코더 및 상기 디코더에 입력되어 상기 인코더 및 상기 디코더를 학습시키는 사용자 동작 인식 시스템.

청구항 6

제 5 항에 있어서,

상기 특징 추출부는 상기 복수 개의 신호 중 상기 제 1 신호를 기준 신호로 설정하고, 상기 제 2 신호 각각에 대하여 상기 제 1 신호를 기준으로 하여 나타내는 상기 특성 정보를 추출하는 것을 특징으로 하는 사용자 동작 인식 시스템.

청구항 7

제 6 항에 있어서,

상기 특성 정보는 상기 제 1 신호와 상기 제 2 신호 사이의 신호값의 차이를 나타내는 정보, 파장의 차이를 나타내는 정보, 파형의 차이를 나타내는 정보, 및 신호의 선형적 상관관계를 나타내는 정보, 신호의 스펙트럼 간의 선형적 상관관계를 나타내는 정보, 정보 흐름의 인과관계를 나타내는 정보, 및 파형의 부호 차이를 나타내는 정보 중 적어도 어느 하나 이상을 포함하는 것을 특징으로 하는 사용자 동작 인식 시스템.

발명의 설명

기술 분야

본 발명은 사용자 동작 인식에 관한 것으로서, 더욱 상세하게는, 신호 간의 관계성을 이용하는 사용자 동작 인

식 방법 및 이를 위한 시스템에 관한 것이다.

배경 기술

- [0002] 최근 정보통신 기술(Information and Communication Technology: ICT)이 고도화되면서, ICT 인프라스트럭처를 통한 데이터 수집 및 인공 지능을 접목하는 지능 정보 기술이 크게 발전하고 있다. 또한, 시장에서 모바일 디바이스 기반의 구글 카드보드(Google Cardboard), 개인용 컴퓨터 기반의 오클러스 리프트(Oculus Rift)와 같은 다양한 플랫폼 상에서 호스팅될 수 있는 실감형 콘텐츠에 대한 수요와 관심이 빠르게 늘어나면서, 사용자의 동작(gesture)을 인식하기 위한 다양한 입력 인터페이스의 연구와 개발이 활발히 진행되고 있다.
- [0003] 종래의 동작 인식 시스템의 일 예는 마이크로소프트 사에 의해 개발된 키넥트(Kinect)로서, 이는 움직임 감지 입력 디바이스를 포함하는 움직임 제어기이다. 키넥트와 같은 움직임 제어기는 RGB 카메라 및 깊이 감지 적외선 카메라와 같은 카메라를 구비한 시각 기반 동작 인식(vision-based gesture recognition)을 수행하며, 사용자가 별도의 동작 인식 디바이스를 부착할 것을 필요로 하지 않는다.
- [0004] 또한, 키넥트 기술의 다양한 애플리케이션에서 사용자의 움직임 정보를 획득하기 위한 소프트웨어 키트가 제공된다. 그러나, 이러한 카메라를 이용한 동작 인식은 벽이나 장애물에 의해 방해받을 수 있다. 또한, 사용자 행동의 반경이 카메라의 방향에 따라 제한적일 수 있는 등 동작을 정확하게 인식하는데 어려움이 있다.
- [0005] 동작 인식 시스템의 다른 예에서는 다축 가속도계와 같은 가속도 센서를 이용하여 동작을 인식하는 시스템들이 있다. 그러나, 이러한 가속도 센서를 이용하는 동작 인식 시스템들은 움직임의 측정이 정확하지 아니할 수 있고, 가속도 센서에 대한 의존성이 너무 높아 정확한 동작의 인지가 어려울 수 있다. 또한, 연속적인 동작 신호 사이의 관계에 대한 고려가 없기 때문에 동적인 동작에 대한 인식의 정확성이 낮을 수 있다.

발명의 내용

해결하려는 과제

- [0006] 본 발명이 해결하고자 하는 과제는 동작 센서를 통해 입력되는 입력 신호 사이의 관계성을 고려함으로써 정확하게 사용자의 동작을 인식할 수 있는 방법을 제공하는 것이다.
- [0007] 본 발명이 해결하고자 하는 다른 과제는 웨어러블 디바이스를 이용하여 사용자의 동작 인식 성능을 향상시킬 수 있는 시스템을 제공하는 것이다.

과제의 해결 수단

- [0008] 상기 과제를 해결하기 위한 본 발명의 일 실시예에 따른 사용자의 동작을 인식하는 인코딩 방법은, 복수 개의 신호, 복원 손실 정보, 및 분류 손실 정보를 수신하는 단계; 상기 복원 손실 정보 및 상기 분류 손실 정보를 이용하여 상기 복수 개의 신호에 대한 특성 정보를 추출하는 단계; 상기 특성 정보를 이용하여 상기 신호의 분류 인덱스를 예측하는 단계; 상기 예측된 분류 인덱스와 오리지널 분류 인덱스 사이의 분류 손실 정보를 산출하는 단계; 및 상기 특성 정보를 출력하는 단계를 포함한다.
- [0009] 또한, 상기 특성 정보를 추출하는 단계는, 상기 복수 개의 신호 중 하나의 신호인 제 1 신호를 기준 신호로 설정하는 단계; 및 상기 제 1 신호를 제외한 제 2 신호 각각에 대하여 상기 제 1 신호를 기준으로 하여 나타내는 상기 특성 정보를 추출하는 단계를 포함한다.
- [0010] 일 실시예에서, 상기 특성 정보는 상기 제 1 신호와 상기 제 2 신호 사이의 신호값의 차이를 나타내는 정보, 파장의 차이를 나타내는 정보, 및 파형의 차이를 나타내는 정보, 신호의 선형적 상관관계를 나타내는 정보, 신호의 스펙트럼 간의 선형적 상관관계를 나타내는 정보, 정보 흐름의 인과관계를 나타내는 정보, 및 파형의 부호 차이를 나타내는 정보 중 적어도 어느 하나 이상을 포함할 수 있다.
- [0011] 상기 과제를 해결하기 위한 본 발명의 다른 실시예에 따른 사용자의 동작을 인식하는 디코딩 방법은, 제 1 신호, 상기 제 1 신호와 상이한 제 2 신호 사이의 관계를 나타내는 특성 정보, 및 복원 손실 정보를 수신하는 단계; 상기 특성 정보, 상기 복원 손실 정보, 및 상기 제 1 신호를 이용하여 상기 제 2 신호를 복원하는 단계; 상기 복원된 제 2 신호와 오리지널 제 2 신호 사이의 비교하여 복원 손실 정보를 산출하는 단계; 및 상기 복원 손실 정보를 출력하는 단계를 포함할 수 있다.
- [0012] 상기 과제를 해결하기 위한 본 발명의 또 다른 실시예에 따른 사용자 동작 인식 시스템은, 복수 개의 신호를 입

력받는 신호 입력부; 복원 손실 정보, 분류 손실 정보, 및 상기 복수 개의 신호를 수신하여 상기 복수 개의 신호 중 제 1 신호와 제 2 신호와의 관계를 나타내는 특성 정보를 추출하는 특징 추출부; 및 상기 복원 손실 정보를 디코더로부터 수신하는 복원 손실 입력부를 포함하는 인코더; 상기 특성 정보를 기초로 하여 분류 인덱스를 예측하는 신호 분류부; 및 상기 예측된 분류 인덱스와 오리지널 분류 인덱스의 분류 손실을 산출하는 분류 손실 산출부를 포함하는 분류기; 및 상기 인코더로부터 상기 특성 정보 및 상기 제 1 신호를 수신하는 신호 수신부; 상기 제 1 신호 및 상기 특성 정보를 이용하여 상기 제 2 신호를 복원하는 신호 복원부; 및 상기 복원된 제 2 신호와 오리지널 제 2 신호를 비교하여 복원 손실을 산출하는 복원 손실 산출부를 포함하는 디코더를 포함하고, 상기 분류 손실 정보는 상기 인코더 및 상기 분류기에 입력되어 상기 인코더 및 상기 분류기를 학습시키고, 상기 복원 손실 정보는 상기 인코더 및 상기 디코더에 입력되어 상기 인코더 및 상기 디코더를 학습시킬 수 있다.

발명의 효과

[0013] 본 발명의 실시예에 따르면, 입력 신호에 대한 가정을 고려하지 아니하고, 하나의 동작을 인식하기 위한 적어도 두 개 이상의 온바디 센서로부터 측정된 입력 신호 데이터만을 이용하여 입력 신호를 분석함으로써, 정확하게 사용자의 동작을 나타내는 신호를 분류할 수 있고, 하나의 입력 신호 자체에 대한 특성 정보가 아닌 복수 개의 입력 신호들의 상관관계를 나타내는 특성 정보를 추출하여 신호를 분류하고, 디코더에서는 상기 특성 정보를 이용하여 수신된 신호 이외의 신호를 복원함으로써 정확하게 사용자의 동작과 관련된 신호를 분류할 수 있는 사용자 동작 인식 방법을 제공하는 것이다.

[0014] 본 발명이 해결하고자 하는 다른 과제는 디코더에서 복원된 신호 및 오리지널 신호 사이의 손실에 대한 복원 손실 정보를 디코더 뿐만 아니라 인코더에서도 이용하여 학습하고, 예측 분류 인덱스와 오리지널 분류 인덱스 사이의 분류 손실 정보 또한 분류기 뿐만 아니라 인코더에서 다시 입력받아 학습에 이용함으로써, 입력 신호 사이의 복잡한 상관관계를 분석할 수 있고, 분석 목적에 맞는 사용자의 동작을 분류한 결과를 획득할 수 있는 사용자 동작 인식 시스템을 제공하는 것이다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 일 실시예에 따른 사용자 동작 인식을 위한 인코더 및 분류기를 나타내는 블록도이다.
 도 2는 본 발명의 일 실시예에 따른 사용자 동작 인식의 인코딩 방법을 나타내는 순서도이다.
 도 3는 본 발명의 일 실시예에 따른 사용자 동작 인식을 위한 디코더를 나타내는 블록도이다.
 도 4는 본 발명의 일 실시예에 따른 사용자 동작 인식의 디코딩 방법을 나타내는 순서도이다.
 도 5는 본 발명의 다른 실시예에 따른 사용자 동작 인식 시스템을 나타내는 블록도이다.
 도 6은 본 발명의 일 실시예에 따른 사용자 동작 인식 방법에 의해 사용자의 동작을 인식한 경우에 대한 정확도 및 일반적인 동작 인식 방법을 이용하여 사용자의 동작을 인식한 경우에 대한 정확도를 비교한 결과이다.

발명을 실시하기 위한 구체적인 내용

[0016] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예를 상세히 설명하기로 한다.

[0017] 본 발명의 실시예들은 당해 기술 분야에서 통상의 지식을 가진 자에게 본 발명을 더욱 완전하게 설명하기 위하여 제공되는 것이며, 하기 실시예는 여러 가지 다른 형태로 변형될 수 있으며, 본 발명의 범위가 하기 실시예에 한정되는 것은 아니다. 오히려, 이들 실시예는 본 개시를 더욱 충실하고 완전하게 하고, 당업자에게 본 발명의 사상을 완전하게 전달하기 위하여 제공되는 것이다.

[0018] 이하, 본 발명의 실시예들은 본 발명의 이상적인 실시예들을 개략적으로 도시하는 도면들을 참조하여 설명된다. 도면들에 있어서, 예를 들면, 부재들의 크기와 형상은 설명의 편의와 명확성을 위하여 과장될 수 있으며, 실제 구현시, 도시된 형상의 변형들이 예상될 수 있다. 따라서, 본 발명의 실시예는 본 명세서에 도시된 부재 또는 영역의 특정 형상에 제한된 것으로 해석되어서는 아니 된다.

[0019] 본 발명의 실시예에 따른 사용자 동작 인식 방법, 장치 및 시스템은 딥러닝을 이용하여 학습할 수 있다. 여기서, 딥러닝(Deep Learning) 기술은, 컴퓨터가 사람처럼 생각하고 배울 수 있도록 하는 인공지능(AI) 기술로서, 인공지능명 이론을 기반으로 복잡한 비선형 문제를 기계가 스스로 학습하여 해결할 수 있도록 한다. 딥러닝은 인간의 두뇌가 수많은 데이터 속에서 패턴을 발견한 뒤 사물을 구분하는 정보처리 방식을 모방해 컴퓨터가 사물

을 분별하도록 기계를 학습시키는 기술이다. 딥러닝 기술을 적용하면 사람이 모든 판단 기준을 정해주지 않아도 컴퓨터가 스스로 인지추론할 수 있게 되고, 음성이미지 인식과 사진 분석 등에 광범위하게 활용하는 것이 가능하다.

- [0020] 즉, 딥러닝은 여러 비선형 변환기법의 조합을 통해 높은 수준의 추상화(abstractions, 다량의 데이터나 복잡한 자료들 속에서 핵심적인 내용 또는 기능을 요약하는 작업)을 시도하는 기계학습(machine learning) 알고리즘의 집합으로 정의되며, 큰 틀에서 사람의 사고방식을 컴퓨터에게 가르치는 기계학습의 한 분야라고 이야기할 수 있다.
- [0021] 어떠한 데이터가 있을 때 이를 컴퓨터가 알아들을 수 있는 형태(예를 들어, 이미지의 경우는 픽셀정보를 열벡터로 표현하는 등)로 표현(representation)하고 이를 학습에 적용하기 위해 많은 연구가 진행되고 있다. 또한, 이러한 노력의 결과로 deep neural networks, convolutional deep neural networks, deep belief networks와 같은 다양한 딥러닝 기법들이 컴퓨터 비전, 음성인식, 자연어처리, 음성/신호처리 등의 분야에 적용되어 최첨단의 결과들을 보여주고 있다.
- [0022] 특히, 컴퓨터 비전의 주요 분야인 영상 인식 및 사물 인식 분야에서의 딥러닝 기술의 중요성이 대두되고 있는데, 자동 음성인식 분야의 자동 음성 번역 및 이해 분야로의 확장과 마찬가지로, 이미지 분류 분야는 자동 영상 캡셔닝(captioning)이라는 더욱 도전적인 분야로 확장되고 있다. 자동 영상 캡셔닝은 딥러닝을 핵심 기반 기술로 사용하는 분야로서, 적용 사례로는 360° 카메라 화면을 이해할 수 있도록 딥러닝을 통해 학습된 자동차 탐재용 컴퓨터 등이 있다.
- [0023] 또한, 컨볼루션 신경망(convolutional neural networks, CNN)은 2차원 영상 처리에 특화되어 인식문제에 주로 사용되어 왔고, 기본적인 핵심 특징으로 첫 번째는 국소 수용 면(local receptive field)을 두어 영상 일부를 표현하는 특징(feature) 추출에 중점을 두었고, 두 번째는 이러한 특징을 나타내는 가중치들을 영상 전 영역에서 공유할 수 있게 하였다는 점에서 막대한 수의 매개변수를 줄이고 영상 위치에 상관없이 특징을 공유할 수 있게 되는 장점이 있다.
- [0024] 세 번째 특징은 복수 개의 계층을 쌓음으로써 상위 계층으로 갈수록 광역 특징을 고려할 수 있다. 최근 딥러닝이 많은 관심을 받게 되면서 convolutional RBM(Restricted Boltzmann Machine)을 계층화한 convolutional deep belief networks, deep CNN과 같이 CNN을 deep 구조화하는 연구들이 진행되어 왔으며 영상 내 객체 인식 및 분류 문제에서 뛰어난 성능을 보이고 있다.
- [0025] 이와 같이, 본 발명의 실시예는 다양한 환경적인 문제로 인한 학습 정보의 질이 떨어지거나 입력 신호에 대한 가정으로 인하여 정확한 학습이 불가능한 문제점을 해결하기 위하여, 동일한 동작을 측정하기 위하여 사용자의 몸 여러 부분에 부착되어 상기 동작시 서로 다른 몸의 부분에서 온바디 센서를 이용하여 측정된 두 개 이상의 실질적인 입력 신호 raw data 만을 이용하여 특성 정보를 추출하고, 상기 특성 정보에 기반하여 신호를 분류함으로써, 종래의 방법에 비해 복잡한 상관관계도 밝혀낼 수 있고 목적에 맞는 사용자의 동작을 인식할 수 있는 사용자 동작 인식 방법 및 이를 위한 시스템을 제공할 수 있다.
- [0026] 도 1은 본 발명의 일 실시예에 따른 사용자 동작 인식을 위한 인코더 및 분류기를 나타내는 블록도이다.
- [0027] 도 1을 참조하면, 사용자의 동작 인식을 위한 인코더 및 분류기는 신호 입력부(10), 특징 추출부(20), 신호 분류부(30), 분류 손실 산출부(40), 및 복원 손실 입력부(50)를 포함할 수 있다. 신호 입력부(10)는 복수 개의 신호를 입력 받을 수 있다. 상기 복수 개의 신호는 신체에 부착된 온바디 센서로부터 측정되는 신호일 수 있고, 상기 센서는 가속도 센서, 거리 센서, 및 이미지 센서와 같은 동작을 인식하기 위한 센서일 수 있다. 또한, 상기 센서로부터 측정되는 신호는 가속도 정보, 거리 정보, 및 영상 정보일 수 있으며, 상기 측정되는 신호의 측정부위 및 갯수는 사용자의 동작 인식의 목적에 맞게 결정될 수 있다. 일 실시예에서, 상기 복수 개의 신호는 하나의 신체로부터 측정되는 가속도 신호일 수 있다.
- [0028] 특징 추출부(20)는 상기 복수 개의 신호를 수신하여 상기 복수 개의 신호 중 제 1 신호와 제 2 신호 사이의 관계를 나타내는 특성 정보를 추출할 수 있다. 상기 제 2 신호는 적어도 하나 이상의 신호일 수 있다. 일 실시예에서, 상기 특성 정보는 상기 제 1 신호와 상기 제 2 신호 사이의 신호값의 차이를 나타내는 정보, 파장의 차이를 나타내는 정보, 파형의 차이를 나타내는 정보, 신호의 선형적 상관관계를 나타내는 정보, 신호의 스펙트럼 간의 선형적 상관관계를 나타내는 정보, 정보 흐름의 인과관계를 나타내는 정보, 및 파형의 부호 차이를 나타내는 정보 중 적어도 어느 하나 이상을 포함할 수 있다. 예를 들어, 상기 특성 정보는 제 1 신호에 대한 제 2 신호값의 비율을 나타내는 정보일 수 있다.

- [0029] 특징 추출부(20)에서 특성 정보가 산출되면 상기 산출된 특성 정보는 상기 제 1 신호와 독자적으로 출력될 수 있다. 이 경우, 특징 추출부(20)는 상기 제 2 신호도 출력하지 아니할 수 있다. 또한, 특징 추출부(20)는 향후 디코더로부터 수신되는 복원 손실 정보를 입력받을 수 있다.
- [0030] 상기 복원 손실 정보는 디코더에서 복원된 제 2 신호와 오리지널 제 2 신호를 비교하여 획득될 수 있으며, 상기 복원 손실 정보를 이용하여 특징 추출부(20)는 특성 정보 산출에 있어 딥러닝을 수행할 수 있다. 특징 추출부(20)는 이후 설명하는 분류 손실 산출부(35)에서 산출되는 분류 손실 정보도 이용하여 특성 정보를 산출할 수 있다. 특징 추출부(20)는 이와 같이 상기 특성 정보를 산출하는데 있어 상기 복원 손실 정보 및 상기 분류 손실 정보를 이용하여 딥러닝을 수행함으로써 정확한 사용자의 동작을 인식하는 효과를 제공할 수 있다.
- [0031] 특징 추출부(20)에서 추출된 특성 정보는 신호 분류부(30)로 입력될 수 있다. 신호 분류부(30)는 입력되는 특성 정보들을 이용하여 $N \times N$ 행렬을 구성할 수 있으며, 예를 들어, 16×16 또는 32×32 의 행렬일 수 있으나 이에 한정되지는 아니한다. 일 실시예에서는, 특성 정보들로 구성되는 $N \times N$ 행렬은 신호 분류부(30)에 입력되기 전에 구성될 수도 있다.
- [0032] 또한, 신호 분류부(30)는 특성 정보를 기초로 하여 입력 신호를 분류할 수 있다. 상기 특성 정보는 입력 신호가 두 개 이상의 신호이므로 두 개 이상의 신호 사이의 관련성을 설명할 수 있으며, 입력 신호들을 특징을 설명할 수 있는 정보일 수 있다. 신호 분류부(30)는 기 입력된 특성 정보로부터 획득한 데이터들을 기초로 하여 상기 입력 신호를 타겟에 맞춰 적절한 클래스로 분류하는 분류 인덱스를 예측할 수 있다. 상기 예측된 분류 인덱스는 기존에 입력되어 학습된 데이터를 이용하므로 학습량이 축적되면서 정확한 예측을 수행할 수 있다.
- [0033] 이후, 예측된 분류 인덱스는 분류 손실 산출부(35)로 입력되어 분류 손실 정보를 생성할 수 있다. 분류 손실 산출부(35)는 예측된 분류 인덱스와 오리지널 분류 인덱스를 비교함으로써 상기 분류 손실 정보를 계산할 수 있다. 분류 손실 산출부(35)에서 산출되는 상기 분류 손실 정보는 이후 신호 분류부(30)에 입력되는 특성 정보를 이용하여 분류 인덱스를 예측하기 위하여 학습에 이용되며, 특징 추출부(20)로도 입력되어 특성 정보 산출을 위한 학습에 이용될 수 있다.
- [0034] 일 실시예에서, 분류 손실 산출부(35)에서 생성되는 상기 분류 손실 정보는 신호 분류부(30)로 재입력될 수 있다. 본 발명의 일 실시예에 따른 사용자 동작 인식 시스템은 상기 분류 손실 정보 뿐만 아니라 이후에 설명되는 복원 손실 정보를 이용하여 딥러닝을 수행하므로 적합한 목적에 맞는 결과를 정확하게 획득할 수 있으며, 사용자의 동작을 정확하게 인식할 수 있다. 신호 분류부(30)는 분류 손실 산출부(35)로부터 입력되는 상기 분류 손실 정보를 이용하여 학습을 수행하기 때문에 사용자의 동작을 정확하게 인식하고 적합한 목적에 맞는 결과를 정확하게 획득할 수 있다.
- [0035] 또한, 복원 손실 입력부(40)는 디코더로부터 산출된 복원 손실 정보를 수신할 수 있다. 상기 복원 손실 정보에는 도 3 및 도 4를 참조하여 디코딩 프로세스에서 설명하기로 한다. 복원 손실 입력부(40)로 입력된 상기 복원 손실 정보는 특징 추출부(20)로 입력될 수 있다. 특징 추출부(20)는 디코더로부터 입력된 이전 특성 정보로부터 복원된 제 2 신호를 이용하여 산출된 상기 복원 손실 정보를 이용한 학습을 통하여 새로이 입력된 입력 신호들로부터 특성 정보를 추출할 수 있다.
- [0036] 일 실시예에서, 상기 복원 손실 정보는 인코더에 입력된 신호 자체에 대한 정보가 아닌 입력된 복수 개의 신호에 대한 복원 손실 정보일 수 있다. 그러므로, 본 발명의 특징 추출부는 복수 개의 입력 신호 데이터로부터 동작 특성 정보를 추출할 수 있으므로, 종래의 입력 신호 자체에 대한 복원 정보를 이용하기 때문에 센서 신호간의 관계성을 고려하지 아니하고 입력 신호만을 분석하여 인식되는 사용자 동작 인식 방법보다 정확하게 사용자의 동작을 인식하고 인식된 동작을 분류할 수 있다.
- [0037] 도 2는 본 발명의 일 실시예에 따른 사용자 동작 인식의 인코딩 방법은 나타내는 순서도이다. 도 2의 사용자 동작 인식 방법은 도 1의 인식 시스템에서 사용자의 동작을 인식하는 방법을 설명하며, 도 1를 참조하여 상술한 설명은 하기에서 도 2를 참조하여 사용자의 동작 인식 방법을 설명시에는 생략될 수 있다.
- [0038] 도 2를 참조하면, 본 발명의 일 실시예에 따른 사용자의 동작 인식 방법은 복수 개의 신호 및 복원 손실 정보를 입력할 수 있다(S10). 상기 복수 개의 신호는 적어도 두 개 이상일 수 있고, 가속도 센서에서 측정된 가속도 신호일 수 있으나 본 발명은 이에 한정되지는 아니한다. 또한, 상기 복원 손실 정보는 이전 입력 신호들에 대한 정보일 수 있다. 일 실시예에서, 상기 복원 손실 정보는 이전 입력 신호들을 복원한 후 생성된 것일 수 있으며, 디코더로부터 입력될 수 있다.

- [0039] 이후, 상기 복수 개의 입력 신호 및 상기 복원 손실 정보를 이용하여 상기 복수 개의 입력 신호들 사이의 관계를 나타내는 특성 정보를 추출할 수 있다(S20). 상기 특성 정보는 상기 복수 개의 입력 신호 중 기준 신호로서 제 1 신호를 설정하고, 상기 제 1 신호와 상기 제 1 신호와 상이한 제 2 신호 사이의 관계를 나타내는 것일 수 있다. 예를 들면, 상기 특성 정보는 상기 제 1 신호와 상기 제 2 신호 사이의 신호값의 차이를 나타내는 정보, 파장의 차이를 나타내는 정보, 파형의 차이를 나타내는 정보, 신호의 선형적 상관관계를 나타내는 정보, 신호의 스펙트럼 간의 선형적 상관관계를 나타내는 정보, 정보 흐름의 인과관계를 나타내는 정보, 및 파형의 부호 차이를 나타내는 정보중 적어도 어느 하나 이상을 포함할 수 있다.
- [0040] 상기 특성 정보를 기초로 하여 입력 신호에 대한 분류 인덱스를 예측할 수 있다(S40). 만일 상기 입력 신호로 하나의 신호만이 입력되는 경우에는 상기 분류 인덱스의 예측은 상기 입력 신호 자체에 대한 것일 수 있다. 그러나, 본 발명의 일 실시예에 따른 사용자 동작 인식 방법은 상기 입력 신호로 상기 제 1 신호 및 상기 제 2 신호를 포함하는 복수 개의 신호가 입력되며, 이 경우 상기 특성 정보는 상기 제 2 신호 자체를 대신하여 상기 제 2 신호와 상기 제 1 신호와의 관계를 나타내는 정보이므로 상기 제 1 신호 뿐만 아니라 상기 제 2 신호에 대한 특성이 반영되는 분류 인덱스를 예측할 수 있다.
- [0041] 이후, 상기 예측 분류 인덱스와 오리지널 분류 인덱스를 비교하여 분류 손실 정보를 산출할 수 있다(S45). 상기 분류 손실 정보는 상기 예측 분류 인덱스와 상기 오리지널 분류 인덱스에 대하여 크로스 엔트로피(cross entropy)를 수행함으로써 획득할 수 있다. 이후, 상기 분류 손실 정보는 다시 상기 분류 인덱스를 예측하는 단계(S40)로 입력됨과 동시에 상기 특성 정보를 산출하는 단계(S20)에도 입력되어 분류 손실이 최소화되어 정확한 분류 인덱스를 예측하고 특성 정보를 산출하는 방향으로 학습될 수 있다.
- [0042] S20 단계에서 추출된 상기 특성 정보는 입력 신호 중 제 1 신호와 함께 디코더로 전송될 수 있다(S30). 상술한 바와 같이, 상기 특성 정보는 입력 신호 그 자체에 대한 특징을 나타내는 정보가 아닐 수 있다. 일 실시예에서, 상기 특성 정보는 디코더로 전송되는 상기 제 1 신호가 아닌 제 2 신호와 상기 제 1 신호와의 관계를 나타내는 특징으로 구성될 수 있다. 그러므로, 복수 개의 입력 신호에 대한 특징이 모두 반영된 특성 정보를 추출하여 학습에 이용함으로써 입력 신호에 대한 가정을 고려함없이 정확하게 입력 신호에 대한 분석이 가능할 수 있다. 또한, 상기 특성 정보만을 디코더로 전송함으로써, 학습 및 분석의 정확성을 향상시킴과 동시에 전송량을 감소시키는 효과를 제공할 수 있다.
- [0043] 도 3는 본 발명의 일 실시예에 따른 사용자 동작 인식을 위한 디코더를 나타내는 블록도이고, 도 4는 본 발명의 일 실시예에 따른 사용자 동작 인식의 디코딩 방법을 나타내는 순서도이다.
- [0044] 도 3 및 도 4를 함께 참조하면, 디코더는 신호 수신부(40), 신호 복원부(50), 및 복원 손실 산출부(60)를 포함한다.
- [0045] 신호 수신부(40)는 인코더로부터 특성 정보 및 외부로부터 입력되는 제 1 신호를 수신할 수 있다(S50). 상기 제 1 신호는 인코더에 입력된 복수 개의 신호 중 하나의 신호일 수 있으며, 센서로부터 측정된 가속도 신호일 수 있다.
- [0046] 상기 제 1 신호는 인코더에 입력된 복수 개의 입력 신호 중 하나의 신호 일 수 있다. 또한, 상기 특성 정보는 상기 제 1 신호 자체에 대한 정보가 아니며, 인코더로부터 입력되지 아니한 제 2 신호와 상기 제 1 신호와의 관계를 나타내는 정보일 수 있다. 예를 들면, 상기 특성 정보는 상기 제 1 신호와 상기 제 2 신호 사이의 신호값의 차이를 나타내는 정보, 파장의 차이를 나타내는 정보, 파형의 차이를 나타내는 정보, 및 신호의 선형적 상관관계를 나타내는 정보, 신호의 스펙트럼 간의 선형적 상관관계를 나타내는 정보, 정보 흐름의 인과관계를 나타내는 정보, 및 파형의 부호 차이를 나타내는 정보 중 적어도 어느 하나 이상을 포함할 수 있다.
- [0047] 이후, 수신된 상기 제 1 신호 및 상기 특성 정보는 신호 복원부(50)로 입력되어 제 2 신호를 복원할 수 있다(S60). 상기 특성 정보는 상기 제 2 신호와 상기 제 1 신호와의 관계를 나타내는 정보이므로 상기 제 1 신호에 적용하여 상기 제 2 신호를 복원할 수 있다. 인코더에 입력된 복수 개의 입력 신호가 세 개 이상으로 쌓을 이루고 있는 경우에는 상기 특성 정보는 상기 제 1 신호를 제외한 2 개의 제 2 신호 각각에 대하여 산출될 수 있고, 상기 제 2 신호인 2 개 신호에 대하여 하나의 정보로 산출될 수 있다.
- [0048] 상기 제 2 신호가 복원되면 복원 손실 산출부(60)는 상기 복원된 제 2 신호와 오리지널 제 2 신호 사이의 차이를 나타내는 복원 손실 정보를 생성할 수 있다(S70). 상기 복원 손실 정보는 신호 복원부(50)로 재입력되어 학습에 이용되며, 학습된 디코더는 이후에 새로이 입력되는 특성 정보를 이용하여 제 2 신호를 더 정확하게 복원할 수 있다. 또한, 상기 복원 손실 정보는 인코더로 송신되어 특정 추출부(20)로 입력될 수 있다(S80).

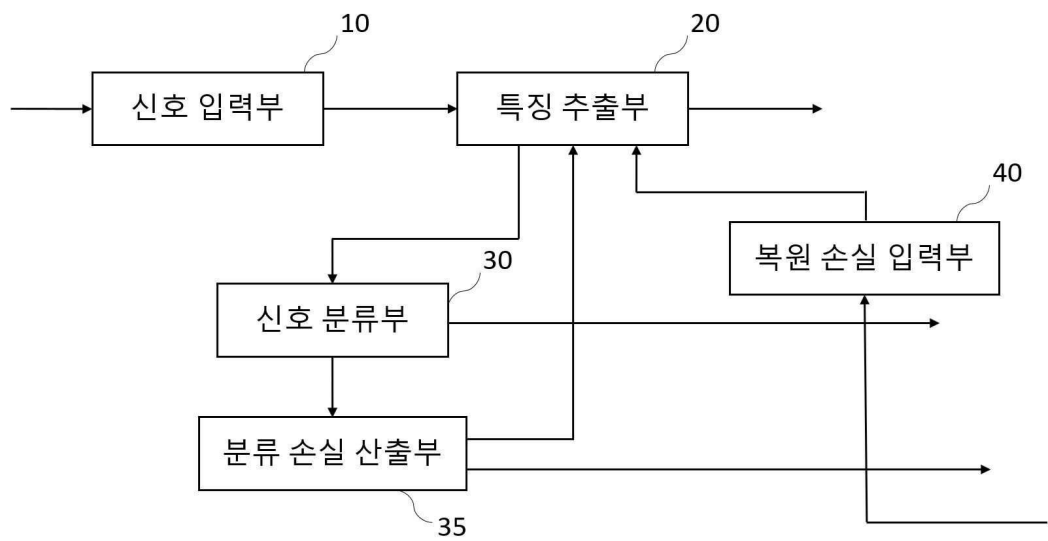
- [0049] 도 5는 본 발명의 다른 실시예에 따른 사용자 동작 인식 시스템을 나타내는 블록도이다.
- [0050] 도 5를 참조하면, 사용자 동작 인식 시스템(1000)은 인코더(100), 분류기(200), 및 디코더(300)를 포함할 수 있다. 인코더(100)는 도 1 및 도 2를 참조하여 설명한 신호 입력부(10) 및 특징 추출부(20)를 포함할 수 있다. 인코더(100)는 제 1 신호(x_i) 및 제 2 신호(x_j)를 포함하는 복수 개의 신호(x_i, x_j)를 입력받아, 제 1 신호(x_i)와 제 2 신호(x_j) 사이의 관계를 나타내는 특성 정보($r_{i \rightarrow j}$)를 생성할 수 있다. 또한, 인코더(100)는 특성 정보($r_{i \rightarrow j}$)를 출력하며, 제 1 신호(x_i) 및 제 2 신호(x_j)는 출력하지 아니할 수 있다.
- [0051] 분류기(200)는 상기 특성 정보($r_{i \rightarrow j}$)를 이용하여 상기 입력 신호에 대한 분류 인덱스(\hat{y}_x)를 예측할 수 있다. 상기 특성 정보($r_{i \rightarrow j}$)는 종래에는 하나의 입력 신호 그 자체에 대한 특징을 나타내는 것과 달리 복수 개의 입력 신호들 사이의 관계를 나타내는 특징에 대한 것일 수 있다. 분류기(200)는 이러한 상기 특성 정보($r_{i \rightarrow j}$)를 이용하여 입력 신호를 분류함으로써 입력 신호에 대한 가정을 배제하고 실질적인 입력 신호로부터 획득되는 정보만을 이용하여 상기 입력 신호들 사이의 관련성을 고려하여 사용자 동작과 관련된 신호를 분류할 수 있다.
- [0052] 또한, 분류기(200)는 예측된 분류 인덱스와 오리지널 분류 인덱스 사이의 차이를 계산하여 분류 손실 정보($\text{loss}_{\text{cls}} = \text{Crossentropy}(\mathbf{y}_x, \hat{\mathbf{y}}_x)$)를 생성할 수 있다. 상기 분류 손실 정보($\text{loss}_{\text{cls}} = \text{Crossentropy}(\mathbf{y}_x, \hat{\mathbf{y}}_x)$)는 분류기(200)를 학습하는데 이용되며, 따라서 이후 입력되는 특성 정보($r_{i \rightarrow j}$)를 정확하게 분류하여 분류 손실 정보($\text{loss}_{\text{cls}} = \text{Crossentropy}(\mathbf{y}_x, \hat{\mathbf{y}}_x)$)가 근소값에 수렴할 수 있도록 도움을 줄 수 있다. 또한, 상기 분류 손실 정보($\text{loss}_{\text{cls}} = \text{Crossentropy}(\mathbf{y}_x, \hat{\mathbf{y}}_x)$)는 분류기(200)를 학습할 뿐만 아니라, 인코더(100)로 송신되어 인코더(100)의 학습에도 이용될 수 있다.
- [0053] 디코더(300)는 인코더(100)로부터 제 1 신호(x_i) 및 특성 정보($r_{i \rightarrow j}$)를 입력받을 수 있으며, 상기 특성 정보($r_{i \rightarrow j}$)를 제 1 신호(x_i)에 적용하여 제 2 신호(\hat{x}_j)를 복원할 수 있다. 또한, 복원된 제 2 신호(\hat{x}_j)와 오리지널 제 2 신호(x_j)사이의 손실을 나타내는 복원 손실 정보($\text{loss}_{\text{rec}} = \text{RMSE}(\mathbf{x}_j, \hat{\mathbf{x}}_j)$)를 산출하여 학습에 이용할 수 있다. 상기 복원 손실 정보($\text{loss}_{\text{rec}} = \text{RMSE}(\mathbf{x}_j, \hat{\mathbf{x}}_j)$)는 인코더(100)로도 입력되어 인코더(100) 내의 특징 추출부(20)를 학습시킬 수 있다. 따라서, 본 발명의 사용자 동작 인식 시스템은 센서로부터 측정된 복수 개의 신호 사이의 관계성을 고려하여 사용자의 동작을 인식할 수 있고, 다른 가정 없이 입력 신호 데이터들 만으로부터 사용자의 동작 인식에 대한 신호를 획득할 수 있으므로 필요에 따라 목적에 맞는 정확한 결과를 획득할 수 있다.
- [0054] 도 6은 본 발명의 사용자 동작 인식 시스템에 의해 측정된 동작 인식의 정확도 및 가중치 F1 스코어(weighted F1 score)(실시예 1) 및 일반적인 사용자 동작 인식 시스템에 의해 측정된 동작 인식의 정확도(비교예 1 내지 4)를 나타낸 것이다. 비교예 1은 K-최근접 이웃(K-nearest neighbor) 알고리즘을 이용하고, 비교예 2는 랜덤 포레스트(random forest) 알고리즘을 이용하며, 비교예 3은 2016년 F.J.Ordonez 및 D.Roggen이 제안한 DeepConvLSTM 알고리즘을 이용하고, 비교예 4는 2018년 Y.Zhao 및 R.Yang 등이 제안한 ResBidirLSTM 알고리즘을 이용하여 사용자의 동작을 측정하였다.
- [0055] 상세하게는, 사용자가 일상적인 행동을 하는 동안, 서로 다른 동작 인식을 위한 복수 개의 가속도 센서를 착용한 채로 stand, walk, sit, lie, 그리고 그 외의 분류되지 않은 상태(null) 중 어느 동작 중에 있었는지를 구분하고자 하였다. 이 경우, 정확도(accuracy)는 (맞게 분류된 동작의 개수)/(전체 동작 개수)로 산출하였다. 또한, 가중치 F1 스코어의 경우, 예를 들어, 가령 stand에 대한 $\text{precision} = (\text{stand 중 맞게 분류된 동작 개수}) / (\text{stand라고 분류된 동작 개수})$, $\text{recall} = (\text{stand 중 맞게 분류된 동작 개수}) / (\text{실제 stand 동작 개수})$ 일 때, stand에 대한 F1 스코어를 $2(\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$ 와 같이 계산하고, 가중치 F1 스코어는 상술한 바와 같이 계산한 각 클래스 별 F1 스코어를 클래스의 크기를 가지고 가중치 합을 하여 산출하였다.
- [0056] 예를 들어, stand, walk, sit, lie, null에 해당하는 동작이 각각 100, 200, 300, 400, 500개인 경우, 가중치 F1 스코어는 $(100 \times (\text{stand의 F1 스코어}) + 200 \times (\text{walk의 F1 스코어}) + 300 \times (\text{sit의 F1 스코어}) + 400 \times (\text{lie의 F1 스코어}) + 500 \times (\text{null의 F1 스코어})) / (100+200+300+400+500)$ 이다.
- [0057] 도 6에 나타난 바와 같이, 본 발명의 일 실시예에 따른 사용자 동작 인식 방법을 이용하여 사용자의 동작을 측

정한 경우(실시예)의 정확도는 0.821으로 일반적인 비교예 1 내지 4의 사용자 동작 인식 방법을 이용하여 동작을 측정한 경우의 정확도인 0.359 내지 0.734 보다 정확도가 높음을 알 수 있다. 또한, 가중치 F1 스코어 측면에서도 실시예에 따른 사용자의 동작 인식시 가중치 F1 스코어(0.814)가 비교예 1 내지 4의 방법을 이용한 가중치 F1 스코어(0.365 내지 0.737) 보다 높음을 확인할 수 있다. 그러므로, 본 발명의 일 실시예에 따른 사용자 동작 인식 방법은 하나의 동작을 측정한 적어도 두 개 이상의 측정 입력 신호의 상관관계를 나타내는 특성 정보를 이용하여 사용자의 동작을 측정하고 분류하고, 복원 손실 정보 및 분류 손실 정보를 인코더에서 재입력받아 학습에 이용함으로써 정확하게 사용자의 동작을 인식할 수 있다.

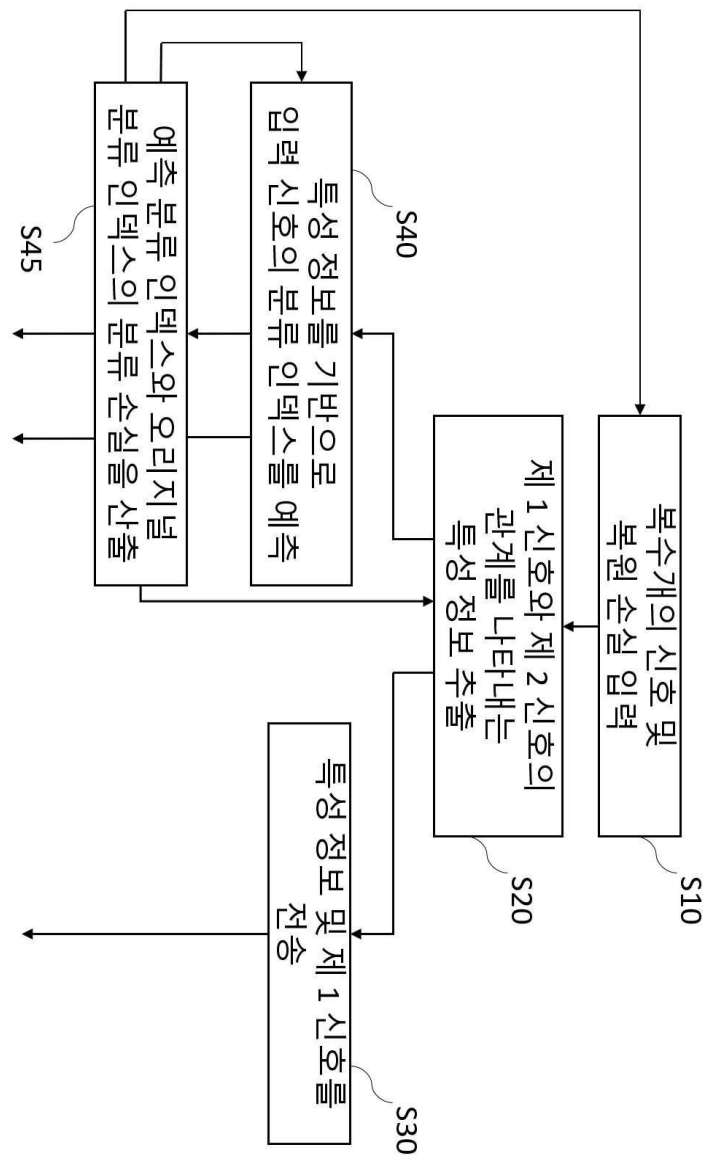
[0058] 이상에서 설명한 본 발명이 전술한 실시예 및 첨부된 도면에 한정되지 않으며, 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러가지 치환, 변형 및 변경이 가능하다는 것은, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 있어 명백할 것이다.

도면

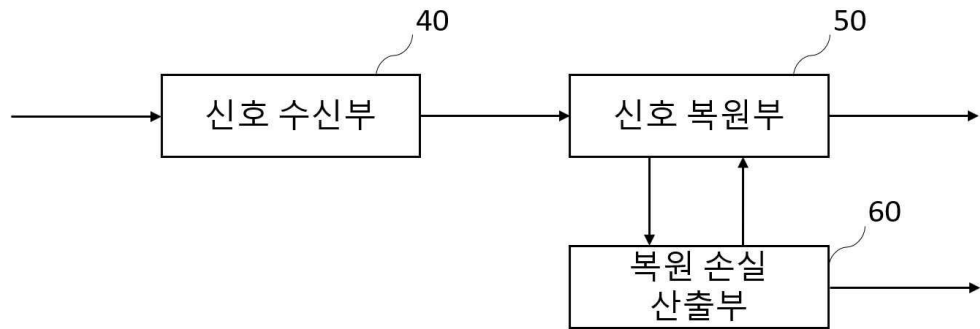
도면1



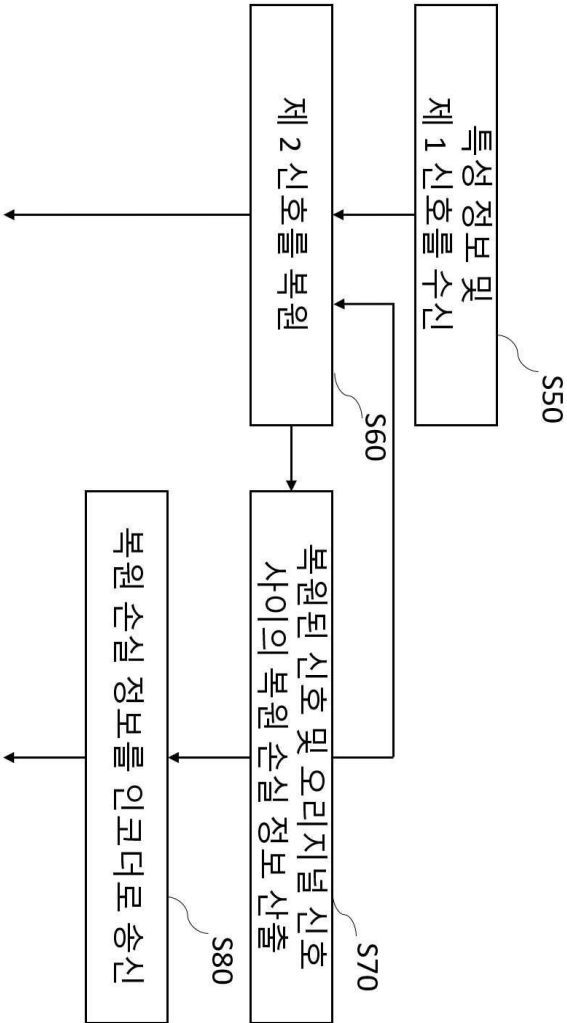
도면2



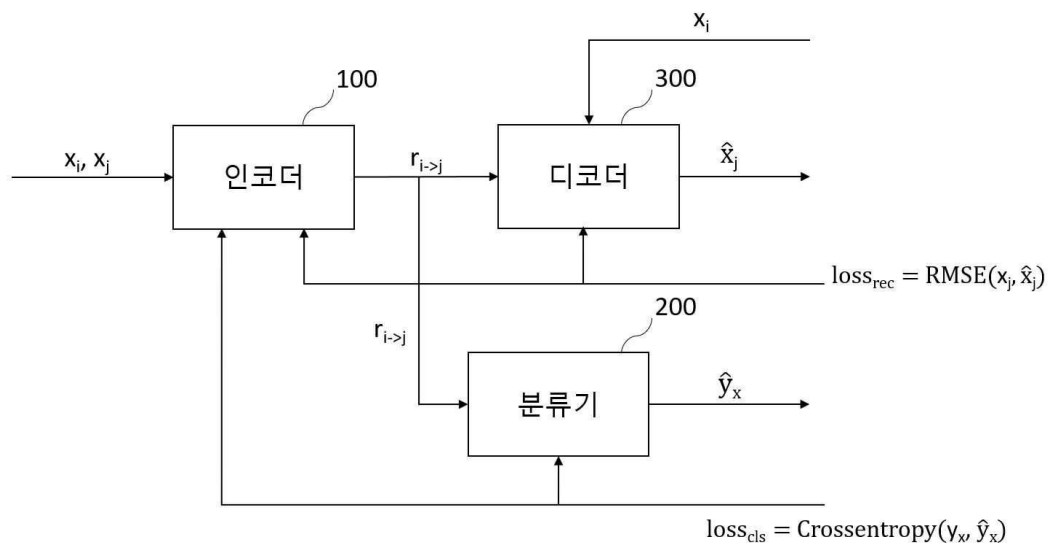
도면3



도면4



도면5



도면6

비교예 및 실시예	정확도	Weighted F1 score
비교예 1(K-nearest neighbor)	0.359	0.365
비교예 2(Random forest)	0.538	-
비교예 3(DeepConvLSTM)	0.734	0.737
비교예 4(ResBidirLSTM)	0.699	0.685
실시예	0.821	0.814